The Heterogeneous Effects of Changing SAT Requirements in Admissions: An Equilibrium Evaluation*

Emilio Borghesan

November 7, 2022

Many universities are reducing emphasis on standardized exam scores in admissions out of concern that the exams limit college access for students from disadvantaged backgrounds. This paper analyzes how such a policy change would affect enrollment patterns and graduation rates at four-year colleges in the United States. To do so, I build an equilibrium model in which colleges rebalance their admissions criteria towards other measurements of students' human capital in the absence of standardized exam scores. The model allows high school students' application decisions and human capital investments to respond endogenously to the admissions policy, while colleges adjust admissions thresholds to maximize their objectives. I estimate the model using data from the Education Longitudinal Study of 2002. I find that banning the SAT causes a small increase in college attendance for low-income students but has a negligible effect on the enrollment of under-represented minority (URM) students, despite estimating that many universities have substantial preferences for diversity. The reason for this result is that endogenous human capital investment and equilibrium responses by capacity-constrained colleges completely offset the diversifying effects of relying more on grades and allowing non SAT-takers to apply to college. Elite colleges are worse off after banning the SAT, as they enroll students with lower skills and see graduation rates drop by 3 pp, while completion rates rise at less selective schools. A separate policy that requires all students to take the exam raises college completion for URMs by 1.6 pp overall by helping schools to identify stronger students.

*Email: borghesa@sas.upenn.edu. I am deeply indebted to Petra Todd, Andrew Shephard, and Francesco Agostinelli for their continued support and mentorship. Special thanks also to Hanming Fang, Michael Gilraine, Margaux Luflade, Aviv Nevo, Holger Sieg, and seminar participants at the University of Pennsylvania.

1 Introduction

In 1935, Harvard University began requiring that all applicants take the Scholastic Aptitude Test (SAT) to be considered for admission. Over the ensuing decades, other colleges followed suit until taking the SAT became a near necessity for high school students interested in attending a four-year college in the United States.[1] But in the last twenty years, this trend has reversed. Between 2001 and 2018, the percentage of four-year public and private non-profit colleges requiring an exam score fell from eighty-five to sixty-seven percent amid concern that the SAT may serve as a barrier to college entry. An extensive literature supports this interpretation by showing how exam-taking mandates, financial incentives, and the opening of testing centers raise college attendance (Klasik 2013, Bulman 2015, Pallais 2015, Hyman 2017, Goodman 2016). Eliminating SAT requirements, however, has not been much studied despite its potential to significantly alter applications, attendance, sorting across schools, and even college completion.

The goal of this paper is to analyze how reducing emphasis on the SAT in admissions would affect patterns of sorting to college and rates of college completion, with particular emphasis on the outcomes for under-represented minorities (URMs) and low-income students.[2] I specify a model of the objectives of admissions departments and how they use the SAT to achieve them. Within the model, eliminating the SAT generates incentives for application behavior and human capital investment among high school students. These behavioral responses may alter the composition of college applicants and subsequently induce capacity-constrained colleges to modify their admissions criteria. This paper takes seriously the notion that students respond to changes in admissions criteria, and that colleges will then have to respond to students' behavior. In such an environment, the theoretical effect of eliminating the SAT on patterns of college attendance and college completion is ambiguous.

Colleges in the model aim to enroll students that are knowledgeable and racially diverse, and they use grades and SAT scores as signals of each student's knowledge at the time of application. Knowledge is treated as a dynamic latent factor that evolves throughout high school in response to inputs that are both exogenous, like family and school characteristics, and endogenous (study time). Eliminating the SAT has two immediate effects. It causes colleges to rely more on the rest of each student's application, their grades and initial conditions at the start of high school, when inferring their knowledge. And, it allows students who have not taken the SAT to apply to college.

[1] Throughout this paper, I use SAT to refer jointly to the SAT and ACT exams.

[2] Under-represented minorities include individuals who identify as Black, Hispanic, Native American, or mixed race.

High school students respond to these two immediate effects by altering how much they study and where they apply to college. The direction of incentives for non SAT-takers is clear. Eliminating the SAT removes a barrier to college entry, raising their incentive to study and their probability of applying to college. However, former SAT-takers may face a reduced incentive to study if grades are not a sufficiently precise signal of their knowledge. The overall effect on patterns of college attendance depends on how these endogenous student responses affect the distribution of knowledge among applicants.

I estimate the model using the Education Longitudinal Study of 2002 (ELS 2002), a rich longitudinal survey of a cohort of students as they transition from high school to college. The ELS 2002 contains extensive information on high school grades, SAT scores, college applications, admissions decisions, and college attendance. I combine the ELS 2002 with data on SAT testing locations and dates, first used in Bulman (2015), and a comparable data set I gathered for the ACT. Consistent with prior literature, I find that greater access to the SAT raises college applications. Distance to college provides an additional source of variation that affects both whether and where to apply to college. Variation in SAT access and distance to college serve as exclusion restrictions that help to identify college preferences. I estimate the model by maximum likelihood using a nested fixed point algorithm. I then use the estimated model to evaluate several counterfactual admissions policies.

I find that eliminating the SAT from consideration at all schools has a negligible effect on the enrollment of URM students and causes a small increase in the enrollment of low-income students. After eliminating the SAT, 30.2% of low-income students enroll in college, compared with 28.9% in the status quo. The policy, however, significantly reduces sorting by knowledge: The average knowledge of students attending elite private colleges falls by 0.20 sd, while it increases by 0.06 to 0.08 sd at the least selective schools. The reduction in assortative matching causes college completion at elite private colleges to fall by 3 pp and to rise at less selective universities.

These key results conflate the effects of the model's four key mechanisms, and an instructive pattern emerges when examining the contribution of each component in isolation. The four mechanisms are a change in admissions criteria when the SAT is eliminated, endogenous applications, endogenous human capital investment, and college optimization in equilibrium. Holding fixed the distribution of applications and test scores in the data, eliminating the SAT muddies application signals and results in less assortative matching by knowledge and household income, which is correlated with knowledge. Allowing for endogenous applications while holding studying constant further reduces sorting and raises enrollment among URMs, who are disproportionately unlikely to take the SAT.

However, allowing for endogenous study time increases stratification by income and knowledge. I estimate that grades are noisy measures of knowledge, so eliminating the SAT reduces the incentive to study among former SAT-takers. Poorer students, who are estimated to have a greater cost of studying, reduce their studying by more and narrowly miss the cut at elite colleges, while richer students remain. In equilibrium, colleges respond to the increase in applications by raising admissions standards and rejecting applicants who would have been marginal admits in partial equilibrium, further reducing college access for low-income and URM students.

Model simulations show that banning the SAT does not raise college attendance for URMs, because there are too few who do not take the exam and who could out-compete those already applying to college in the status quo. Banning the SAT causes applications from URMs to increase by 30% versus 17% for white and Asian students, but the new applicants, who are on average less-skilled than the status quo applicants, mostly fail to gain admission as the surge in applications bids up admissions thresholds. I show how a hypothetical intervention that could raise skills at the start of high school for URMs who do not take the SAT would instead enable many of them to out-compete SAT-takers, causing an increase in both college attendance and college completion for URM students.

My approach to modeling admissions is a significant departure from the prior literature, which typically models admission to college as a threshold crossing model in terms of a continuous index of observed variables (for example, Kapor 2020). Such a method falters when some of the measurements comprising the index, say SAT scores, are no longer used. By microfounding admissions criteria as a search for a dynamically evolving latent factor using whatever measurements colleges observe under a specific policy regime, my approach causes admissions offices to automatically rebalance their criteria towards grades when SAT scores are gone. This method delivers a probability of admission for every student, with or without an SAT score.

I modify the standard dynamic factor model of Cunha, Heckman, and Schennach (2010) and Agostinelli and Wiswall (2020) by including demographic-specific measurement parameters and by letting the initial distribution of knowledge vary by a set of covariates – such as mother's education, income, and race – that are likely correlated with investment prior to high school.[3] Together with the use of threshold rules for ad-

[3]Typically, dynamic factor models that estimate the technology of skill formation for students are estimated on young children, and an assumption of equal starting conditions is reasonable (Cunha, Heckman, and Schennach 2010, Agostinelli and Wiswall 2020). However, the model in this paper begins in the ninth grade, and a history of unequal educational investments will result in students commencing the ninth grade with different skills. Allowing for differences in the initial distribution of knowledge turns out to have large

mission that vary by demographic, these modifications reproduce a unique feature of the college market in the United States, namely that admissions offices interpret grades and test scores relative to each student's background.[4] Within this framework, the same grade or SAT score will cause admissions offices to more aggressively update their prior if the measurement is particularly informative for that student or if observable factors put that student at an initial disadvantage. Admissions departments in my model use the Kalman Filter, which corrects for any biases present in grades and SAT scores, to generate these updates. Given the intense scrutiny of admissions practices by parents, the media, and the courts, it seems reasonable to assume that admissions officers conduct some sort of filtering process when forming educated guesses of each student's academic preparation.[5]

Estimates of the dynamic factor model reveal that studying is productive: An increase of ten hours per week causes knowledge to rise by 0.08–0.10 sd each year. I also find that much has been decided by the start of high school. The difference in ninth grade knowledge between students whose mother has a graduate degree and those whose mothers are high school dropouts is 0.90 sd. URMs begin high school at a 0.65 sd disadvantage relative to white and Asian students. I do not find any evidence of bias in the SAT math or verbal exams. If anything, GPAs show more evidence of bias than standardized tests. Moreover, GPAs are noiser than the SAT, and they are noisier for URMs than for white and Asian students. This suggests that colleges will struggle to identify highly skilled URM candidates for admission if they must rely more on grades.

In light of the inability of the No-SAT policy to enroll additional URM students, I investigate an alternative policy, recommended in Dynarski (2018), that mandates all high school students take the SAT. Like the No-SAT counterfactual, this policy removes a barrier to college application, but it does so without reducing the amount of information available to colleges. Relative to the No-SAT policy, SAT-for-All causes the fraction of URMs attending a four-year college to rise by 1.4 pp and college completion to increase by 1.6 pp as colleges manage to identify more skilled students for admission. Low-income student enrollment is similar to the No-SAT policy. I also evaluate an SAT-Optional policy at elite colleges and find that it slightly increases URM attendance at elite private colleges.

effects on college attendance and completion in the model.

[4]Interpreting grades and test scores in the context of each student's educational opportunities and family background has been common since at least the 1990s. Bowen and Bok (2016) provide examples of conversations with admissions deans that reveal how the same grades and SAT scores may affect the admissions probabilities of several hypothetical applicants in heterogeneous ways.

[5]The Massachusetts Institute of Technology explains that they use a combination of factor analysis and thresholds to determine whether a student qualifies for admission. A blog on their website explains their recent decision to reinstate the SAT, stating "... we do not consider an applicant's [SAT] scores at all beyond the point where preparedness has been established as part of a multifactor analysis." (Schmill 2022)

This paper shows that modeling general equilibrium is important in analyzing the effects of large changes in college admissions. Papers that estimate equilibrium models of the market for college admissions in the United States include Epple, Romano, and Sieg (2006, 2008) and Fu (2014). The paper with the closest modeling approach to this one is Kapor (2020), which analyzes the Texas Top Ten Percent Plan (TTP), a policy that granted all students graduating in the top ten percent of their high school class acceptance to any Texas public university. My paper shares the three-part application-admission-matriculation equilibrium of Kapor (2020) and Fu (2014), but I add several novel features. I allow the distribution of grades and test scores observed by colleges to be endogenous with respect to the policy, I microfound college preferences for student characteristics, and I analyze the effects of admissions policies on college completion.[6]

The endogenous mechanisms in the model are motivated by a growing literature that shows how pre-college human capital investment responds to changes in admissions policies. Tincani, Kosse, and Miglino (2021) and Cotton, Hickman, and Price (2022) conduct experiments to show that high school students trade off leisure and the probability of admission to college when deciding how much to invest in their skills. Leeds, McFarlin, and Daugherty (2017), Golightly (2019), and Akhtari, Bau, and Laliberté (2020) exploit changes in admissions policies at Texas public universities to show that effort increases when policies render admission more likely but decreases when admission becomes certain. Grau (2018) and Bodoh-Creed and Hickman (2017) similarly find that effort in high school is shaped by admissions criteria in Chile and the US. Bond et al. (2018) and Goodman, Gurantz, and Smith (2020) show how applications respond endogenously to SAT scores. This paper incorporates these multiple mechanisms in an equilibrium framework and demonstrates the quantitative importance of each channel.

This paper is organized as follows. The next section describes the data used for the analysis and presents summary statistics. Section 3 describes the model that is taken to the data and discusses some of its properties. Sections 4 and 5 discuss identification and estimation of the model. Section 6 shows the estimated model parameters, while section 7 presents estimates of counterfactual policies and explains the mechanisms behind the results. Section 8 concludes.

[6]Several empirical papers analyze how affirmative action affects labor market outcomes and whether it could lead to mismatch between students and colleges (Otero, Barahona, and Dobbin 2021, Arcidiacono et al. 2011, Arcidiacono 2005). Dillon and Smith (2017) consider whether uncertainty in the admissions process leads to mismatch. Arcidiacono, Kinsler and Ransom (2022a, 2022b) show how preferences at Harvard University vary by race.

2 Data

This study uses data from the Education Longitudinal Study of 2002 (ELS 2002). The ELS 2002 randomly samples a nationally representative cohort of students who were in the tenth grade in 2002 and follows them through high school, college, and into the labor market. Students are surveyed four times, in 2002, 2004, 2006, and 2012. In 2006, students are either in college or participating in the labor market and receive surveys tailored to their status. The 2012 survey wave, eight years after graduation from high school, records educational attainment.

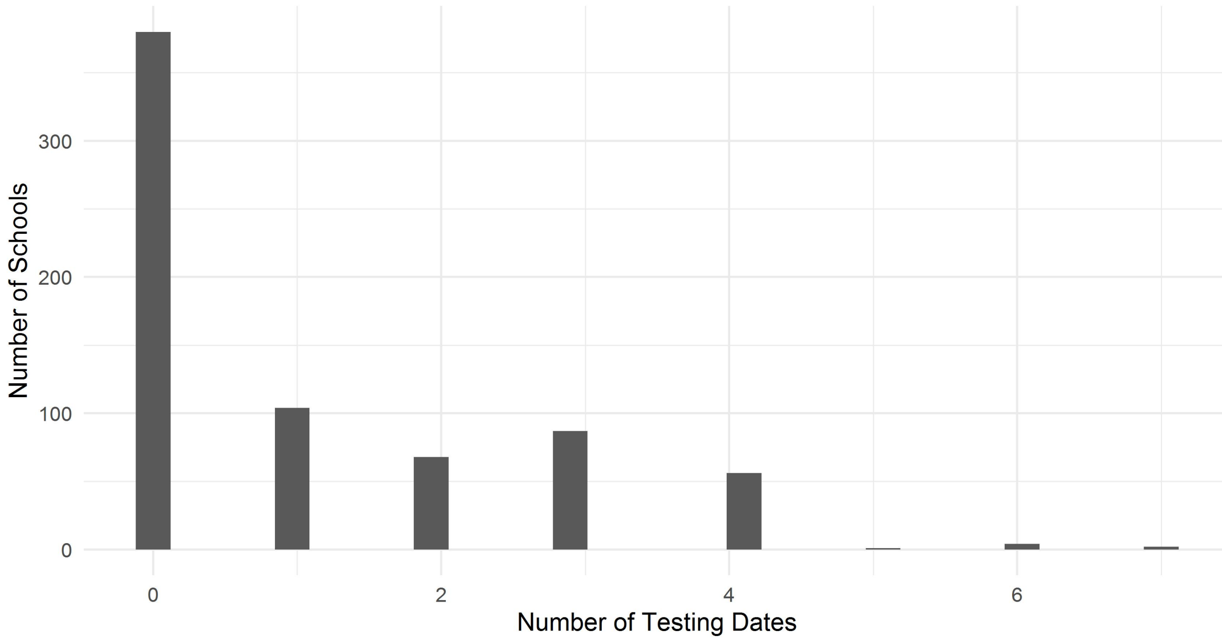
The ELS 2002 contains multiple measurements of cognitive skills throughout high school. Grade-point averages (GPAs) for each year of high school have been converted to a common scale and are weighted by Carnegie units.[7] The ELS 2002 also contains SAT scores in math and verbal skills obtained from the College Board, and ACT scores in math, English, reading, and science obtained from ACT, Inc. In addition, the National Center for Education Statistics (NCES) administers exams in math and reading to all students in the ELS 2002 in grades ten and twelve. These multiple measurements across time make it possible to estimate a dynamic factor model of skill formation for students in high school.

I merge additional sources of data into the ELS 2002. I combine a database of SAT testing center dates first used in Bulman (2015) with information from yearly ACT test registration booklets I obtained from ACT, Inc. to construct a measure of exam access. I define access to be the number of testing dates at one's own high school during spring of the junior year, when students typically take the SAT. Figure 1 plots the distribution of exam access across students in the sample. The modal number of testing dates per school is zero, the mean is 1.09 days, and some schools that host both the SAT and ACT exams have up to seven testing dates during the course of the semester.

Test centers open after an employee at a particular school, typically a teacher or guidance counselor, volunteers to act as a test coordinator and applies to the College Board or ACT, Inc. to host the exam on a specific day. Testing sites must satisfy certain criteria, like having a quiet examination room and a secure location to store materials, but most applications are approved. The key factor in a school becoming a testing center is therefore whether someone at that school takes the initiative to apply. Bulman (2015) surveys fifty testing coordinators to understand their motivations. Many expressed concern that nearby testing centers were at capacity and a desire to offer their students the exam in a familiar environment. In this paper, I control for school type (private, public, Catholic),

[7]A Carnegie unit corresponds to one course taken every day, for one period per day, for a full school year.

Figure 1: SAT and ACT Testing Dates per School



The figure shows the number of dates reserved for SAT or ACT examinations during the spring of 2003 at schools in the ELS 2002. SOURCES: College Board, ACT, Inc. and U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

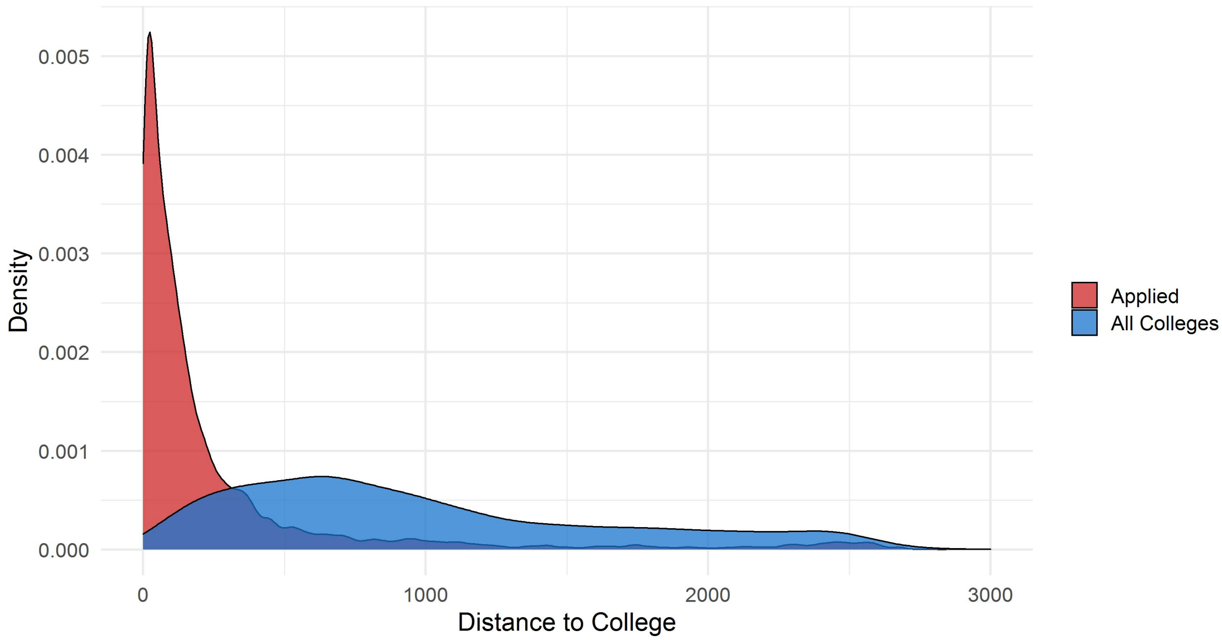
total enrollment, and poverty rates within the school district, all of which are likely to influence demand for testing facilities. I argue that residual variation in SAT access is quasi-random, stemming in part from differences across schools in whether an employee decides to apply.

This paper also exploits variation in distance to college to shift college applications. Figure 2 shows that, relative to distance to all colleges, the distribution of distance to colleges where students apply is heavily weighted towards zero, suggesting that distance may shape application decisions. Carneiro and Heckman (2002) and Cameron and Taber (2004) express concern that distance to school may be correlated with student ability and thus endogenous. The model that I describe in the next section addresses this concern by letting distance shift demand for college conditional on a precise measure of ability, a student's posterior knowledge after grades and test scores in each year of high school are revealed.

After removing observations with missing data, the sample has $9,910$ observations.[8] Tables 1 and 2 present summary statistics for this sample. The measurements, standardized to have zero mean and unit variance, indicate that URMs have lower standardized

[8] Appendix A provides details regarding the construction of the sample.

Figure 2: Distance to College



The figure shows the density of distance to college for students in the ELS 2002. Distance is computed between the centroid of each student's home census block in the ELS 2002 and the latitude and longitude of each college in IPEDs. The figure plots both the unconditional density of distance in blue and the density of distance to schools applied to in red. Distances above 3,000 miles (relevant only for Alaska and Hawaii) have been truncated. SOURCE: National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

test scores and GPAs. URMs have less educated mothers, are more likely to grow up in a household headed by a single parent (typically a mother), and are more likely to have been retained prior to high school. They also attend schools with higher class sizes, where more students qualify for free or reduced-price lunch, and they grow up in families with an average income that is about $23,000 less than white and Asian families. URMs actually attend schools with more SAT and ACT testing dates in the spring of their junior year, but this is largely due to white and Asian students attending smaller schools and Catholic or private schools where the exams are rarely held. After controlling for school size, location, type, and the poverty of the school district, URM students have lower exam access, as indicated by the variable "Residualized Num Testing Dates."

Table 2 demonstrates that URM students are less likely to take the SAT, less likely to attend college, less likely to complete college, and less likely to complete conditional on attending college. They also study fewer hours while in high school. When URMs attend college, they attend very different colleges than white and Asian students. Figure 3 indicates that URMs are under-represented at highly selective colleges and state flagships, but are over-represented at less selective private universities and public satellite colleges. The

Table 1: Summary Statistics, I

	URM		White & Asian	
	Mean	SD	Mean	SD
Measurements				
GPA, 9th grade	-0.31	0.98	0.22	0.91
GPA, 10th grade	-0.31	0.97	0.22	0.91
NCES Reading, 10th grade	-0.34	0.94	0.22	0.95
NCES Math, 10th grade	-0.40	0.93	0.26	0.92
GPA, 11th grade	-0.32	0.98	0.20	0.92
SAT Math	-0.57	0.90	0.08	0.93
SAT Verbal	-0.51	0.92	0.09	0.95
GPA, 12th grade	-0.35	1.02	0.19	0.92
NCES Math, 12th grade	-0.38	0.92	0.25	0.94
Exogenous Variables				
Student-teacher Ratio	17.40	4.46	16.19	4.13
Free Lunch	0.28	0.23	0.15	0.16
Num Testing Dates	1.06	1.53	0.96	1.44
Residualized Num Testing Dates	-0.02	1.47	0.05	1.35
Household Income	50,000	40,500	72,800	49,000

The table presents summary statistics for knowledge measurements and high school inputs in the ELS 2002. The knowledge measurements have been standardized by their population mean and standard deviation. NCES refers to standardized exams administered to all students as part of the ELS 2002. Free Lunch refers to the percentage of students at the student's school who qualify for a free or reduced-price lunch. Num Testing Dates refers to the number of SAT or ACT testing dates held at a student's school during the spring of their junior year of high school. This number is then residualized on controls for school type, enrollment, and district poverty rates. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

Table 2: Summary Statistics, II

	URM	White & Asian
Choices		
Study Hours, per week	5.87	6.49
Take SAT	0.63	0.79
Attend 4-yr College	0.33	0.49
Complete 4-yr College	0.19	0.34
Complete 4-yr College Given Attendance	0.59	0.70
Initial Conditions		
Female	0.52	0.50
Retained before High School	0.10	0.06
Single Parent	0.32	0.16
Mother : High School	0.25	0.27
Mother : Some College	0.35	0.34
Mother : 4-year Degree	0.14	0.22
Mother : Postgraduate	0.07	0.11
Observations	2860	7050

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

model I describe in the next section will offer an explanation for why URM students are less likely to complete college and evaluate whether alternative admissions mechanisms can raise college completion.

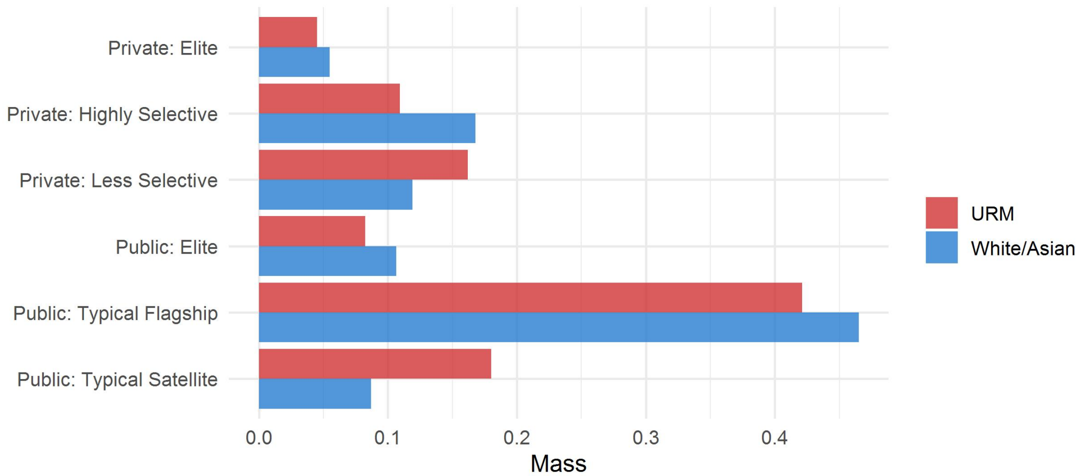
3 Model

In this section, I present the model that I use to analyze how eliminating the SAT affects patterns of college attendance and completion. The model describes high school students' endogenous application and human capital investment behavior and colleges' optimization problem, which is affected by whether they observe the SAT.

3.1 Timing

The model has three time periods: high school, college transition, and college completion. I treat ninth grade GPA as an initial condition, and at the beginning of tenth grade students choose how much time to allocate to studying and whether they will take the SAT.

Figure 3: Type of College Attended Among College-Goers



The figure shows the fraction of all college students of each demographic group that are enrolled at each college tier. The grouping of colleges into tiers is described in section 3.9. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

This choice of study time will take effect for three years. At the end of high school, grades and SAT scores are realized for each student. These measurements depend on a student's grade nine knowledge, as well as time devoted to study, educational inputs, and idiosyncratic shocks. After these measurements are realized, the second period of the model, the transition to college, begins.

Similar to Kapor (2020) and Fu (2014), the transition to college consists of three parts: the application phase, admissions, and matriculation. First students apply to college, and then colleges decide whom to admit conditional on the applications they've received. Once students receive a set of admissions offers, they decide where to matriculate.

College completion, up to eight years later, occurs in the final period of the model. Completion depends on students' knowledge when matriculating to college, the type of college they attend, and exogenous factors including first generation college attendance, household income, and race.

3.2 Technology and Measurement System

Knowledge evolves deterministically as a result of prior knowledge, study time, and educational inputs, $\mathbf{I}_{i,t}$.[9] The technology of skill formation is allowed to differ by whether an

[9]Throughout the paper, I use bold font to denote vectors and matrices.

individual belongs to an under-represented minority to allow for the possibility that differences in the marginal productivity of study time and schooling inputs may influence study decisions. Knowledge evolves according to the following value-added equation:

$$\log K_{i,t} = \gamma^{K,R} \log K_{i,t-1} + \beta^{H,R} h_{i,t} + \mathbf{I}'_{i,t} \boldsymbol{\beta}^{I,R} \ , \tag{1}$$

where $R \in \{URM, WA\}$ denotes parameters that are specific to either URM or white and Asian students.

I model grade point averages (GPAs) and standardized tests as noisy measures of knowledge. In each grade, this mapping is

$$\mathbf{y}^R_{i,t} = \boldsymbol{\mu}^R_t + \boldsymbol{\alpha}^R_t \log K_{i,t} + \boldsymbol{\varepsilon}^R_{i,t} \ , \tag{2}$$

where $\boldsymbol{\mu}^R_t$ is a vector of constants, $\boldsymbol{\alpha}^R_t$ is a vector of factor loadings, and $\boldsymbol{\varepsilon}^R_{i,t}$ is a vector of normally-distributed disturbances with mean zero and a diagonal covariance matrix. As with the technology of skill formation, the measurement system is allowed to vary by URM status in an unrestricted way. This parameterization allows me to conduct inference on whether the SAT is biased against URMs and whether the informativeness of grades varies by demographic groups. I define bias and differential signal informativeness for measurement j in year t as follows:

$$\text{Bias} := \mu^{URM}_{t,j} - \mu^{WA}_{t,j} \ , \tag{3}$$

$$\text{Differential Signal Informativeness} := \frac{\alpha^{URM}_{t,j}}{\sigma^{URM}_{t,j}} - \frac{\alpha^{WA}_{t,j}}{\sigma^{WA}_{t,j}} \ . \tag{4}$$

In section 6, I will conduct inference on bias and differential signal informativeness for each measurement in the ELS 2002. The informativeness of grades, in particular, may affect college completion in a world without the SAT by potentially impeding colleges' ability to select qualified students for admission.

3.3 Initial Conditions

The first observed measure of knowledge for students in the ELS 2002 is in the ninth grade, and it is unlikely that all students begin high school on a level playing field. Provided that the skill accumulation equation in (1) holds in middle and elementary school, we can use backwards substitution to write knowledge in the ninth grade as a function of an entire history of study decisions and educational inputs as follows:

$$\log K_{i,9} = f(h_{i,1}, \dots, h_{i,9}, \mathbf{I}_{i,1}, \dots, \mathbf{I}_{i,9}, K_{i,0}) \ .$$

This equation means that even if children were born with equal endowments, $K_{i,0}$, a history of unequal investments would generate differences in the distribution of $K_{i,9}$. While the ELS 2002 does not record the entire history of inputs and study decisions prior to high school, I *do* allow the distribution of $K_{i,9}$ to vary according to a set of predetermined covariates, $\mathbf{W}_i$, that are likely to be correlated with prior investments. Hence, rather than imposing the normalization that $\log K_{i,9} \sim N(0, \sigma_k^2)$, as is common in the literature estimating dynamic factor models, I instead allow the distribution of initial conditions to vary by $\mathbf{W}_i$ as follows:

$$\begin{pmatrix} \log K_{i,9} \\ \mathbf{y}_{i,9}^R \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{W}_i'\mathbf{a} \\ \boldsymbol{\mu}_9^R + \boldsymbol{\alpha}_9^R \mathbf{W}_i'\mathbf{a} \end{pmatrix}, \begin{pmatrix} \sigma_k^2(\mathbf{W}_i) & \sigma_k^2(\mathbf{W}_i)\boldsymbol{\alpha}_9^{R'} \\ \sigma_k^2(\mathbf{W}_i)\boldsymbol{\alpha}_9^R & \sigma_k^2(\mathbf{W}_i)\boldsymbol{\alpha}_9^R\boldsymbol{\alpha}_9^{R'} + \boldsymbol{\Sigma}_9^R \end{pmatrix} \right) , \tag{5}$$

where the variance of ninth grade knowledge is given by $\sigma_k^2(\mathbf{W}_i) = \exp(\mathbf{W}_i'\mathbf{b})$ and $\boldsymbol{\Sigma}_9^R = \mathbb{E}(\boldsymbol{\varepsilon}_{i,9}^R \boldsymbol{\varepsilon}_{i,9}^{R'})$.

3.4 The Kalman Filter

The model treats grades and test scores as noisy measures of a dynamically evolving latent state. Admissions officers observe these measurements and form expectations over each student's knowledge at the time of application by using the Kalman Filter. Eliminating the SAT affects admissions decisions through changing the set of measurements available to filter this latent state.

Let $URM_i \in \{0, 1\}$ indicate whether a student belongs to an under-represented minority, and define the initial information set by

$$\Omega_{i,9} := \{URM_i, \mathbf{W}_i, \mathbf{y}_{i,9}, \{\mathbf{I}_{i,k}\}_{k=10}^{12}\} \ ,$$

and subsequent updates as $\Omega_{i,t} := \{\Omega_{i,t-1}, \mathbf{y}_{i,t}, h_{i,t}\}$.[10] The Kalman Filter yields the following update for knowledge after observing the initial conditions and ninth grade GPA:

$$\log K_{i,9} \mid \Omega_{i,9} \sim N(m_{i,9}, P_{i,9}) \ ,$$

[10] Individuals have full information over the realization of future educational inputs. The sources of incomplete information in the model are over future actions, realizations of test scores, admissions, and college completion.

where

$$
\begin{aligned}
m_{i,9} &:= \sigma_k^2(\mathbf{W}_i)\boldsymbol{\alpha}_9^{R'}\mathbf{F}_{i,9}^{-1}\left(\mathbf{y}_{i,9} - (\boldsymbol{\mu}_9^R + \boldsymbol{\alpha}_9^R\mathbf{W}_i'\mathbf{a})\right) \ , \\
P_{i,9} &:= \sigma_k^2(\mathbf{W}_i) - \sigma_k^2(\mathbf{W}_i)\boldsymbol{\alpha}_9^{R'}\mathbf{F}_{i,9}^{-1}\boldsymbol{\alpha}_9^R\sigma_k^2(\mathbf{W}_i) \ , \\
F_{i,9} &:= \sigma_k^2(\mathbf{W}_i)\boldsymbol{\alpha}_9^R\boldsymbol{\alpha}_9^{R'} + \boldsymbol{\Sigma}_9^R \ .
\end{aligned}
$$

For an individual with information set $\Omega_{i,t-1}$ who chooses to study $h_{i,t}$ hours, the prediction for period t knowledge is

$$
\begin{aligned}
\log K_{i,t} \mid \Omega_{i,t-1}, h_{i,t} &\sim N(m_{i,t|t-1}, P_{i,t|t-1}) \ , \\
m_{i,t|t-1} &:= \gamma^{K,R} m_{i,t-1} + \beta^{H,R} h_{i,t} + \mathbf{I}_{i,t}'\boldsymbol{\beta}^{I,R} \ , \\
P_{i,t|t-1} &:= \gamma^{K,R^2} P_{i,t-1} \ ,
\end{aligned}
$$

and the subsequent update of knowledge is

$$
\begin{aligned}
\log K_{i,t} \mid \Omega_{i,t} &\sim N(m_{i,t}, P_{i,t}) \ , \\
m_{i,t} &:= P_{i,t|t-1}\boldsymbol{\alpha}_t^{R'}\mathbf{F}_{i,t}^{-1}\left(\mathbf{y}_{i,t} - (\boldsymbol{\mu}_t^R + \boldsymbol{\alpha}_t^R m_{i,t|t-1})\right) \ , \\
P_{i,t} &:= P_{i,t|t-1} - P_{i,t|t-1}\boldsymbol{\alpha}_t^{R'}\mathbf{F}_{i,t}^{-1}\boldsymbol{\alpha}_t^R P_{i,t|t-1} \ , \\
\mathbf{F}_{i,t} &:= \boldsymbol{\alpha}_t^R P_{i,t|t-1}\boldsymbol{\alpha}_t^{R'} + \boldsymbol{\Sigma}_t^R \ .
\end{aligned}
$$

3.5 Preferences

3.5.1 Colleges

Colleges are grouped into tiers, $c = 1, \ldots, C$, each comprising a continuum of capacity-constrained colleges that have preferences over the knowledge and diversity of matriculating students.

An application to college is defined to be a pair, (S_i^c, URM_i), where S_i^c is a scalar signal of student i's knowledge at the time of application and URM_i is the student's demographic. The signal is a draw from the distribution of twelfth grade knowledge conditional on scores observable by each college: $S_i^c \sim f(\log K_{i,12} \mid \Omega_{i,12}) = N(m_{i,12}, P_{i,12})$.[11]

I define an admissions policy to be a mapping from applications to acceptances and

[11] Drawing signals from $f(\log K_{i,12} \mid \Omega_{i,12})$ presumes that colleges observe study effort. A more realistic approach would have colleges integrate over the distribution of study effort, rather than observing it, when deciding whom to admit. It is, however, unlikely that after observing measurements in each year of high school, $y_{i,9}, \ldots, y_{i,12}$, integrating over study time would generate markedly different predictions of $K_{i,12}$. For computational reasons, this approach was not adopted.

rejections:

$$\text{Policy} : \mathbb{R}^N \times \{0,1\}^N \to \{0,1\}^N \ . \tag{6}$$

Define λ^c_{URM} to be the fraction of students matriculating to college c who belong to an under-represented minority: $\lambda^c_{URM} := \mathbb{P}(URM_i = 1 \mid Attend_{i,c} = 1)$. Colleges in tier c have a production function that is increasing in a signal of knowledge and diversity, and they choose an acceptance policy to solve

$$\begin{aligned} \max \quad & \mathbb{E}[\kappa S^c_i + (1-\kappa) \log \lambda^c_{URM}] \\ \text{s.t.} \ & \sum_{i=1}^{N} \mathbb{P}(Attend_{i,c} = 1 \mid \Omega_{i,12}) = N^c \ , \end{aligned} \tag{7}$$

where N^c is the number of students in the ELS 2002 who attend college c. The expectation is taken over the probability that a student matriculates conditional on the admissions policy. This specification assumes that admissions offices maximize a weighted sum of a signal of knowledge and diversity and that they satisfy their capacity constraints in expectation.[12] The weights are allowed to vary by college tier. While an admissions policy is defined to be a mapping from the space of applications to the space of acceptances and rejections, the structure of the problem leads to the following proposition, which states that admissions policies can be characterized by a pair of threshold rules.

Proposition 1. *The optimal policy for college c is a pair of demographic-specific threshold rules: (S^{c*}_0, S^{c*}_1).*

Proof. Let the policy rule for students with $URM_i = 1$ be arbitrary. Suppose the policy rule for students with $URM_i = 0$ is not a threshold rule. Then there exist two students, l and m, with probabilities of matriculation conditional on admission given by p_l and p_m, such that $S^c_l > S^c_m$ but $\mathbb{P}(Accept_l) < 1$ and $\mathbb{P}(Accept_m) > 0$. Consider the following modified admission policy: $\tilde{\mathbb{P}}(Accept_l) = \mathbb{P}(Accept_l) + \varepsilon$, $\tilde{\mathbb{P}}(Accept_m) = \mathbb{P}(Accept_m) - \varepsilon \frac{p_l}{p_m}$. The modified acceptance rule satisfies the constraint and leaves λ^c_{URM} unchanged, but improves the objective function by $\varepsilon\kappa\,(S^c_l - S^c_m)\,p_l > 0$. Hence, the optimal policy for $URM_i = 0$ students is a threshold rule in S^c_i.

Let the policy rule for students with $URM_i = 0$ be arbitrary. By similar argument, the optimal admissions policy for $URM_i = 1$ students is a threshold rule. Hence, the

[12] I do not directly model tuition and assume that both tuition setting and financial aid formulas are invariant to the counterfactuals explored in this paper.

optimal admissions policy is a pair of threshold rules, (S_0^{c*}, S_1^{c*}), that (potentially) differ by demographic. $\square$

The use of threshold rules combined with a preference for URMs implies that colleges would choose to accept only URM students if they had greater average knowledge than white and Asian students and existed in the population in sufficient proportions. This is, however, not an empirically relevant scenario. The distribution of skills in the ELS 2002 together with a value of $\kappa < 1$ will cause the threshold to be lower for URMs than for white and Asian students, $S_1^{c*} < S_0^{c*}$.

3.5.2 Students

Students in the model choose how much to study, whether to take the SAT while in high school, and where to apply to college upon graduation. I first describe the three parts of the college transition phase – application, admission, and matriculation – in reverse chronological order before characterizing the problem of a high school student.

Matriculation: The indirect utility function for a student who attends college c is expressed as the following linear function of a fixed effect for that school, tuition, distance to college, financial aid, the probability of completing college, and a shock that is distributed Type 1 extreme value:

$$U_{i,c}(\Omega_{i,12}) = \underbrace{\bar{U}_c + \beta_T Tuition_{i,c} + \beta_{D,c} Dist_{i,c} + \beta_A Aid_{i,c} + \beta_P P(Complete_{i,c} = 1 \mid \Omega_{i,12}, c)}_{V_{i,c}} + \varepsilon_{i,c} \ . \tag{8}$$

The probability of completing college depends on $\Omega_{i,12}$, meaning that students form their expectation based on measurements that are observed by both students and colleges. Distaste for distance, $\beta_{D,c}$, varies by whether the college is public or private. $V_{i,c}$ represents the deterministic component of utility. I normalize the deterministic value of not attending college to zero, $V_{i,0} = 0$.[13] Students choose from among their admissions portfolio, B, the option that maximizes their utility. The chosen option, C_i, satisfies

$$C_i = \arg\max_{c \in B} \{U_{i,c}\} \ , \tag{9}$$

[13] A future version of the paper, currently being estimated, allows $V_{i,0}$ to vary with local wages.

and the probability of making choice C_i given admissions portfolio B is

$$P(C_i = c \mid B, \Omega_{i,12}) = \frac{\exp(V_{i,c})}{1 + \sum_{k \in B} \exp(V_{i,k})} \ . \tag{10}$$

The value of being admitted to portfolio B is given by the following log-sum term:

$$U_{i,B} := \mathbb{E}[\max_{c \in B} U_{i,c}] = \log\left(1 + \sum_{c \in B} \exp V_{i,c}\right) \ .$$

Admissions: Each student's application portfolio is transformed into an admissions portfolio depending on whether their application signals exceed the thresholds at the schools where they apply. I assume that application signals are iid draws from f($\log K_{i,12} \mid \Omega_{i,12}$), so the probability that student i obtains admissions set B given application set A is

$$P(B \mid A, \Omega_{i,12}) = \prod_{c \in B} \mathbb{P}(S_i^c > S_{URM_i}^{c*} \mid \Omega_{i,12}) \prod_{d \in A \setminus B} \mathbb{P}(S_i^d < S_{URM_i}^{d*} \mid \Omega_{i,12}) \ . \tag{11}$$

The distribution $f(\log K_{i,12} \mid \Omega_{i,12})$ is fully characterized by its mean and variance, $(m_{i,12}, P_{i,12})$, so I replace $P(B \mid A, \Omega_{i,12})$ with $P(B \mid A, m_{i,12}, P_{i,12})$. This means that, regardless of the number of measurements in $\Omega_{i,12}$, the state space for each individual is only two-dimensional. High school students who are deciding how much to study and whether to take the SAT form expectations over $(m_{i,12}, P_{i,12})$ rather over the realization of each individual GPA and exam score.

Application: Applicants to portfolio A pay a fixed cost of applying to each school and then a marginal cost of applying to additional schools within the same tier. Fixed and marginal costs may vary by school, so that the total application cost can be written as

$$cost_i(A) = \sum_{c=1}^{C} FC_{i,c}(A) + MC_{i,c}(A) + \varepsilon_{i,A} \ , \tag{12}$$

where $\varepsilon_{i,A}$ represents unobserved factors that shift application costs and is modeled as a Type 1 Extreme Value shock with scale parameter λ_A. The fixed cost varies with household income, whether the mother has a college degree, and the distance between student i's home while in high school and college c, while the marginal cost is a fraction of the fixed cost. Letting $n_c(A)$ denote the number of applications to school c in portfolio A, I specify the fixed and marginal costs as follows:

$$FC_{i,c}(A) = \mathbb{1}_{c\in A}\left(\delta_c^{(1)} + \delta_c^{(2)} Inc_i + \delta_c^{(3)} MomCollege_i + \delta_c^{(4)} Dist_{i,c}\right) \ ,$$
$$MC_{i,c}(A) = \max\{n_c(A) - 1, 0\}\left(\delta_c^{(5)} FC_{i,c}(A)\right) \ ,$$

where $MomCollege_i = 1$ if individual i's mother has a college degree.

Students who apply to a portfolio, A, obtain a benefit that integrates over the expected utility of every possible admissions subset, B, that could be obtained from A. Formally, the utility of submitting application portfolio A is:

$$V_i^{Coll}(m_{i,12}, P_{i,12}, A) = \sum_{B\in A} P(B \mid A, m_{i,12}, P_{i,12}) U_{i,B}(m_{i,12}, P_{i,12}) - cost_i(A) \ , \tag{13}$$

where $U_{i,B}$ is written as a function of the state variables $(m_{i,12}, P_{i,12})$ to denote that the value of admissions set B depends on the probability of completing college at each school within B, which in turn varies with the mean and variance of the student's knowledge when enrolling in college.

The student's portfolio choice problem is

$$A_i(m_{i,12}, P_{i,12}, SAT_i) = \underset{A\in\mathcal{A}(SAT_i)}{\arg\max}\, \{V_i^{Coll}(m_{i,12}, P_{i,12}, A)\} \ . \tag{14}$$

The set of possible application portfolios, $\mathcal{A}(SAT_i)$, depends on whether the student took the SAT while in high school. $\mathcal{A}(1)$ is the universe of all possible application portfolios, while $\mathcal{A}(0) = \{0\}$, because all four-year colleges required the SAT during this time. I vary $\mathcal{A}(0)$ under counterfactual policy regimes that allow students to apply to college even without an SAT score. The probability of applying to application set A is

$$P(A \mid m_{i,12}, P_{i,12}, SAT_i) = \frac{\exp\left(\frac{V_i^{Coll}(m_{i,12}, P_{i,12}, A)}{\lambda_A}\right)}{\sum_{A'\in\mathcal{A}(SAT_i)} \exp\left(\frac{V_i^{Coll}(m_{i,12}, P_{i,12}, A')}{\lambda_A}\right)} \ . \tag{15}$$

The value of beginning the college application phase with state variables $(m_{i,12}, P_{i,12}, SAT_i)$ is given by the log-sum term

$$\begin{aligned}\overline{V}_i^{Coll}(m_{i,12}, P_{i,12}, SAT_i) &:= \mathbb{E}[max_{A\in\mathcal{A}(SAT_i)} V_i^{Coll}(m_{i,12}, P_{i,12}, A)] \\ &= \lambda^A \log\left(\sum_{A\in\mathcal{A}(SAT_i)} \exp\left(\frac{V_i^{Coll}(m_{i,12}, P_{i,12}, A)}{\lambda_A}\right)\right)\end{aligned}$$

High School: Students in high school have preferences over hours spent studying, taking the SAT, and their expectation of admission to college as follows:

$$U_i^{HS}(a) = (\gamma_H + \gamma_H^{Inc} Inc_i) H_i(a) + (\gamma_S + \gamma_S^Z Z_i^{SAT} + \gamma_S^{Inc} Inc_i) SAT_i(a) + \varepsilon_i(a) + \int \overline{V}_i^{coll}(m_{i,12}, P_{i,12}, SAT_i(a)) dF(m_{i,12}, P_{i,12} \mid \Omega_{i,9}, a) \ , \tag{16}$$

where a refers to the action chosen by the student. The disutility for hours spent studying is allowed to vary by household income, Inc_i, and the preference for taking the SAT is allowed to vary by both income and exam access, Z_i^{SAT}, as defined in section 2. This permits the model to capture logistical challenges that limit students' ability to take the exam and thus apply to college. $H_i(a)$ denotes the average amount of time spent studying each week while in high school. Students make this decision at the beginning of 10th grade, and study time is assumed to take effect for three consecutive years, i.e. $h_{i,t}(a) = H_i(a)$ for $t = 10, 11, 12$ in equation 1. $\varepsilon_i(a)$ is a Type 1 Extreme Value shock.

High school students choose an action to maximize their utility subject to the technology of skill formation and the measurement system. Their problem is written as follows:

$$\max_a U_i^{HS}(a) \tag{17}$$

subject to

$$\log K_{i,t} = \gamma^{K,R} \log K_{i,t-1} + \beta^{H,R} H_i(a) + \mathbf{I}'_{i,t} \beta^{I,R} \ ,$$
$$\mathbf{y}_{i,t}^R = \boldsymbol{\mu}_t^R + \boldsymbol{\alpha}_t^R \log K_{i,t} + \boldsymbol{\varepsilon}_{i,t}^R \quad \text{for t = 10, 11, 12 and } R \in \{URM, WA\} \ .$$

Note that students do not observe $K_{i,t}$. The model assumes that students perform the Kalman Filter to forecast the distribution of $(m_{i,12}, P_{i,12})$ given their initial conditions, $\Omega_{i,9}$, and their choice, a. From the point of view of the student, what matters is how their actions while in high school influence their probability of admission to college and their chance of completing college. The true value of $K_{i,t}$ is irrelevant for admissions decisions, because colleges base their decisions on signals drawn from $f(\log K_{i,12} \mid \Omega_{i,12})$ and $K_{i,12} \notin \Omega_{i,12}$. Regarding college completion, knowing the true value of $K_{i,12}$ might help students predict their chances of completing college at each school in their admissions set. However, for simplicity's sake, I do not add $K_{i,t}$ as an additional state variable, and instead let college completion depend on the same state variables, $(m_{i,12}, P_{i,12})$, as college admission. This implies that students do not precisely know their own cognitive skill but learn about it from grades and SAT scores.

3.6 College Market Equilibrium

A *College Market Equilibrium* is defined as a set of policy functions for students $\{a_i, A_i, C_i\}_{i=1}^N$, a set of threshold rules $\{S_{URM}^{c*}\}_{URM=0}^1$ for colleges $c = 1, \ldots, C$, and a distribution of grades and SAT scores, $\Omega_{i,12}$, such that

1. Given the admissions thresholds $\{S_{URM}^{c*}\}_{URM=0}^1$ for $c = 1, \ldots, C$, and the state variables in each period, the policy functions, $\{a_i, A_i, C_i\}_{i=1}^N$, solve students' maximization problems in 17, 14, and 9; and

2. The admissions thresholds, $\{S_{URM}^{c*}\}_{URM=0}^1$ for $c = 1, \ldots, C$, maximize colleges' objective function in 7 subject to their capacity constraints, taking as given the realized distribution of student test scores, $\{\Omega_{i,12}\}_{i=1}^N$, the applications that have been submitted, $\{A_i\}_{i=1}^N$, and the actions of other colleges; and

3. The distribution of realized scores $\{\Omega_{i,12}\}_{i=1}^N$ that colleges take as given is consistent with the initial conditions, $\{\Omega_{i,9}\}_{i=1}^N$, and the student actions that produce these scores, $\{a_i\}_{i=1}^N$.

The equilibrium notion is a standard Nash Equilibrium, and it is characterized by a system of equations in terms of the best response functions for all C colleges. These best response functions are derived from the first-order conditions of the college optimization problem in equation 7 and form a system of $2 \times C$ equations in $2 \times C$ unknowns. These equations are:

$$\frac{\frac{\partial \mathbb{P}(Attend_{i,c}=1)}{\partial K_c^{1*}}}{\frac{\partial \mathbb{P}(Attend_{i,c}=1)}{\partial K_c^{0*}}} = \frac{\frac{\partial \mathbb{E}[S_i^c|Attend_{i,c}=1]}{\partial K_c^{1*}} - \kappa_c \frac{\partial \mathbb{P}(i=URM|Attend_{i,c}=1)}{\partial K_c^{1*}}(\mathbb{P}(i = URM|Attend_{i,c} = 1) - \lambda^{URM})}{\frac{\partial \mathbb{E}[S_i^c|Attend_{i,c}=1]}{\partial K_c^{0*}} - \kappa_c \frac{\partial \mathbb{P}(i=URM|Attend_{i,c}=1)}{\partial K_c^{0*}}(\mathbb{P}(i = URM|Attend_{i,c} = 1) - \lambda^{URM})},$$

$$\sum_{i=1}^N \mathbb{P}(Attend_{i,c} = 1 \mid \Omega_{i,12}) = N^c ,$$

where λ^{URM} is the population fraction of under-represented minorities.

I do not prove existence or uniqueness of the equilibrium. For certain extreme parameter values, colleges will be unable to satisfy their capacity constraints and an equilibrium will not exist. However, such a scenario is not empirically relevant. I have always been able to solve for an equilibrium, and I have never found multiple equilibria for a fixed set of parameter values. Different starting guesses for the admission thresholds converge to the same equilibrium, and small perturbations of the parameters produce equilibria with nearby thresholds. This suggests that the optimizer is not jumping between equilibria as it searches over the parameter space.

3.7 College completion

The ELS 2002 records whether each individual obtains a bachelors degree within eight years of graduating from high school. I model college completion as a function of $K_{i,12}$ and the tier of school, c, the student attends:

$$Complete_{i,c} = \mathbb{1}(\omega_c^{(1)} + \omega_c^{(2)} \log K_{i,12} + \mathbf{X}_i' \omega^{(3)} + \eta_i > 0) \ , \tag{18}$$

where $\eta_i \sim N(0,1)$. $O_{i,c}$ can be thought of as a production function for degree attainment that depends on student inputs, $K_{i,12}$, college inputs, $\omega_c^{(1)}$, and controls, $\mathbf{X}_i \subseteq \mathbf{W}_i$.[14] I let the constants in equation 18 vary by college, thereby capturing both observed and unobserved factors that influence rates of completion at each college. Students, who do not observe $K_{i,12}$, compute their completion probability by integrating over it using the distribution $f(\log K_{i,12} \mid \Omega_{i,12})$ when deciding where to matriculate. Hence, $P(Complete_{i,c} = 1 \mid \Omega_{i,12})$ in equation 8 is given by

$$P(Complete_{i,c} = 1 \mid \Omega_{i,12}) = \int \mathbb{P}(Complete_{i,c} = 1 \mid \log K_{i,12}) dF(\log K_{i,12} \mid \Omega_{i,12}) \tag{19}$$

3.8 Discussion

The model assumes that each college prioritizes the knowledge and racial diversity of its student body. Specifying college preferences in this way is consistent with their mission statements, nearly all of which express a desire to enroll diverse and academically prepared students.[15] Papers that model college admissions all assume that colleges value cognitive skill, but the literature is divided over what else they value. Epple, Romano, and Sieg (2006) give colleges preferences over socioeconomic diversity, while Kapor (2020) and Epple, Romano, and Sieg (2008) add preferences for racial diversity. My choice of giving preferences for racial diversity is consistent with a principle goal of many who advocate for eliminating the SAT, namely to increase access to college for URMs (Soares 2020). Even without giving schools direct preferences for socioeconomic diversity, the model closely matches sorting to college by household income (section 6.4).

The model has admissions offices forming expectations of each student's knowledge at the time of application using the Kalman Filter. This framework treats the SAT as just

[14] $\mathbf{X}_i$ includes URM status, family income, and mother's education.

[15] Of the top 50 universities currently ranked by US News and World Report, 47 have clearly defined mission statements, of which 42 mention knowledge directly in the statement and 41 mention diversity or have a separate statement affirming a commitment to diversity. Schools that do not directly mention these words use related words like intellectual, discovery, and inclusion.

one of many signals of latent knowledge, albeit one whose bias and informativeness may differ from grades. The assumption that admissions offices are sophisticated, rational agents is strong but not unreasonable. Admissions offices at the most popular colleges typically employ over twenty full-time workers, and hire many others each winter, to read applications. Many schools make decisions only after multiple meetings in which admissions officers make arguments for and against admitting each student (Tough 2019).

Eliminating the SAT in the model sets in motion a range of behavioral responses by high school students. First, it allows a new pool of students to apply to college. The model predicts that those who actually apply will have a low application cost, as determined by income, distance to college, and mother's education, and either a high probability of admission or a high value for college attendance. The estimated model will show that students greatly value a college degree (β_P in equation 8), meaning that the students most likely to apply to college under the new policy are the ones most likely to complete it.

Second, removing the SAT generates contrasting effects on study behavior. Students in the model trade off the cost of studying today with a greater probability of admission to college after high school. The model lets the cost of studying vary by income, which is a flexible way of capturing various factors, like working for pay or having limited space at home to study, that may affect the study decision of low-income teenagers. Students who do not take the test will now be able to apply to college, raising their incentive to study.[16] At the same time, the removal of a potentially informative signal weakens the link between studying and college acceptance. Which of these effects prevails likely depends on the student's place within the initial skill distribution and her cost of effort.[17] The total effect of removing the SAT on college attendance boils down to how these endogenous mechanisms affect the distribution of knowledge among college applicants in the new equilibrium.

There is some debate over whether rational expectations (RE) or some other form of expectations best characterize student perceptions of the admissions process.[18] In this

[16]At the time students in the ELS 2002 were applying to college, over 85% of colleges within each tier required an SAT or ACT score. The model assumes that an exam score is required to apply to college.

[17]The model suggests that one reason URM students may study less than white and Asian students are the barriers to taking the SAT and thus applying to college. When these barriers are removed, URMs may close the study gap with their white and Asian peers.

[18]Cotton, Hickman, and Price (2022) show in a field experiment that investment in human capital in the presence of affirmative action (AA) is consistent with rational expectations. Arcidiacono et al. (2020) collect data on subjective earnings expectations and occupational choice probabilities and find that they are highly predictive of future earnings and occupational choices. On the side of biased expectations, Hastings, Neilson, and Zimmerman (2015) demonstrate that students who choose unprofitable college degree programs considerably overestimate the earnings of past graduates while high-ability students have relatively accurate beliefs. Wiswall and Zafar (2015) and Delavande and Zafar (2019) find that providing accurate information about wages causes students to update their *beliefs* but has little effect on their *choices*, suggesting

paper, I give students RE over their probability of admission to college and over the effect of study time on academic performance. The ELS 2002 does not provide sufficiently rich data on subjective expectations to permit a major departure from RE. Nevertheless, the model does an impressive job of matching patterns in the data (section 6.4).

3.9 Aggregation

This paper analyzes attendance and completion at four-year colleges. I group colleges according to a combination of their Barron's selectivity ranking and type (public vs private non-profit). The exact groupings are depicted in Table 3. These groupings have been chosen so that the analysis can speak to admissions practices at an identifiable set of schools – like elite public and private universities and state flagships – while retaining a sufficiently large sample size in each group to estimate preferences. Tiers one through three correspond to private colleges and universities, ranked in descending order of selectivity, while tiers four through six are public universities, ranked in descending order of selectivity. Community colleges are grouped together with no college as part of the outside option. Colleges in the same tier are assumed to have the same admissions threshold and colleges of the same type are assumed to have the same preferences. The model therefore allows, say, private colleges to have a stronger preference for diversity than public colleges. Classifying colleges in this manner is consistent with the purpose of the Barron's selectivity rankings, which aim to group schools together that have a common admissions standard. In addition, allowing preferences for diversity to differ by type can capture the extent to which state legislatures and the courts have placed limits on the discretion of admissions offices at public universities.

Although it is computationally convenient, aggregation creates challenges. Preferences for college (equation 8) depend on tuition and the distance student i would need to travel to attend school c. Which of the many schools within tier c should should determine the values of tuition and distance, $(Dist_{i,c}, Tuition_{i,c})$? In the analysis that follows, I choose the reference school for individual i to be the closest school within tier c.[19] Tuition for this reference school is in-state tuition when the student and school are located in the same state and out-of-state tuition otherwise.

I allow students to send multiple applications to each college tier. To limit the size of

the presence of large nonpecuniary preferences for the type of college and field of study.

[19] In many cases, this is the very school to which students apply. When students in the ELS 2002 apply to a school within in particular tier, the school they apply to is the closest one to their home between 20% and 50% of the time, depending on the tier. When the school that students are considering is not the closest within a given tier, it will often be the second closest, and distance to the school under consideration will be correlated with distance to the closest school, thereby limiting the severity of measurement error.

Table 3: College Groupings

Type	Tier	Barron's Rank	Description	Examples
Private	1	1	Elite	Harvard, Swarthmore, Northwestern, USC
	2	2/3	Highly selective	University of Miami, DePaul, Pepperdine
	3	4/5/6	Less selective	University of Mobile, Concordia University-St. Paul, Monmouth University
Public	4	1/2	Elite	UCLA, UIUC, Georgia Tech, UNC-Chapel Hill
	5	3	Most state flagships	Wisconsin-Madison, University of Arizona, most SUNY campuses
	6	4/5/6	Satellite campus, some flagships	Alabama A&M, Boise State, Northern Kentucky

the choice set, however, I limit applications to two per tier. Any student sending more than two applications to a given tier is coded as sending exactly two. I do not allow students to send applications to all possible permutations of colleges, but instead limit the application set to the set of unique application portfolios in the data. Hence, while there are $3^6 = 729$ potential portfolios with up to two applications per tier, I allow students in the model to choose from among the 584 unique portfolios observed in the data.

3.10 Financial Aid

The model allows financial aid to shape matriculation choices through equation 8. Any study of college attendance must deal with the fact that aid is unobserved at schools to which the student does not apply. Another challenge is that the ELS 2002 contains data on federal financial aid, but not state or institutional aid.

I address this problem of partially observed data by training a random forest on data from a more recent NCES educational survey, the High School Longitudinal Study of 2009 (HSLS 2009). The HSLS 2009 is a longitudinal survey of the transition from high school to college with many of the same measurements as the ELS 2002 (GPA, SAT scores, college attended). One advantage of the HSLS 2009 over the ELS 2002 is that it records federal, state, and institutional financial aid awarded to each student.[20] I use the HSLS 2009 data

[20]The main disadvantage of the HSLS 2009 relative to the ELS 2002 and the reason it was not used for this study is that it does not record each student's entire admissions and acceptance portfolio.

set to construct a set of predictors that are likely to influence the amount of financial aid received – grades, SATs, family structure, number of siblings, state of residence, college attended – and train a random forest using five-fold cross validation to predict the proportion of tuition covered by financial aid at each college tier. I then construct the same predictors in the ELS 2002 and use the random forest to predict the proportion of tuition that would be covered by financial aid for each individual in the ELS 2002 at each college tier. I work with proportions rather than aid dollars to control for the growth in college tuition between the 2002 and 2009 cohorts. $Aid_{i,c}$ is therefore the predicted proportion of tuition at school c that would be covered by financial aid awarded to individual i.[21]

4 Identification

4.1 Dynamic Factor Model

When students submit an application to college, the admissions office observes a signal drawn from the distribution of latent knowledge conditional on $\Omega_{i,12}$. Estimating the model requires identifying the parameters that determine this distribution, which consist of the skill technology (equation 1) and the measurement system (equations 2 and 5).

Identification of the dynamic factor model follows from arguments in the literature (Cunha, Heckman, and Schennach 2010; Agostinelli and Wiswall 2020; Williams 2020). Because knowledge is a deterministic function of observables, it is possible to write the entire vector of measurements throughout high school as a function of the initial knowledge draw, $\log K_{i,9}$.[22] To reduce notational clutter, the following equations condition on $\boldsymbol{\mu}_{i,t}$, H_i, $\mathbf{W}_i$, and $\mathbf{I}_{i,t}$:

$$\begin{pmatrix} \mathbf{y}_{i,9} \\ \mathbf{y}_{i,10} \\ \mathbf{y}_{i,11} \\ \mathbf{y}_{i,12} \end{pmatrix} = \underbrace{\begin{pmatrix} \boldsymbol{\alpha}^R_{i,9} \\ \gamma^{K,R}\boldsymbol{\alpha}^R_{i,10} \\ \gamma^{K,R^2}\boldsymbol{\alpha}^R_{i,11} \\ \gamma^{K,R^3}\boldsymbol{\alpha}^R_{i,12} \end{pmatrix}}_{\mathbf{A}} \log K_{i,9} + \begin{pmatrix} \boldsymbol{\varepsilon}^R_{i,9} \\ \boldsymbol{\varepsilon}^R_{i,10} \\ \boldsymbol{\varepsilon}^R_{i,11} \\ \boldsymbol{\varepsilon}^R_{i,12} \end{pmatrix},$$

[21] The model gives students perfect foresight over financial aid, $Aid_{i,c}$. If students instead had to form expectations over aid, their applications would likely appear more random and less targeted to their preferred school. An important extension of this paper would have students form expectations over financial aid. More research is needed to understand how the prospect of aid shapes application decisions.

[22] Appendix B explores whether adding a stochastic shock to equation 1 affects the inferences drawn from the dynamic factor model. It does not, and since identifying a model with this shock requires additional normalizations beyond those discussed in this section, the main analysis was conducted using a deterministic skill technology.

where $\log K_{i,9}$ is one-dimensional, and $\mathbf{y}_{i,t}$, $\boldsymbol{\alpha}_t^R$, and $\boldsymbol{\varepsilon}_{i,t}^R$ are vectors with lengths that vary by t. As long as $\mathbf{A}$ contains at least three measurements, $\mathbf{A}$ satisfies the row deletion property and it is possible to separately identify $\mathbf{A}\Phi\mathbf{A}'$ from $\mathbf{\Sigma}_\varepsilon$, where $\Phi := var(\log K_{i,9})$.[23] A further normalization is needed to separately identify $\mathbf{A}$ and Φ. This is achieved by excluding a constant from $\Phi(\mathbf{W}_i) = \exp(\mathbf{W}_i'\mathbf{b})$, so that $\mathbf{W}_i = \mathbf{0}$ implies that $\Phi(\mathbf{W}_i) = 1$. With this normalization, $(\mathbf{A}\Phi\mathbf{A}')_{1,1}$ identifies $\boldsymbol{\alpha}_9$.

$\mathbf{A}$ is now separately identified from Φ, but it is still necessary to identify $\gamma^{K,R}$ separately from the other factor loadings, $\boldsymbol{\alpha}_{i,10}$, $\boldsymbol{\alpha}_{i,11}$, and $\boldsymbol{\alpha}_{i,12}$. It would, in principle, be possible to scale up $\gamma^{K,R}$ by c and scale down $\boldsymbol{\alpha}_{i,10}$, $\boldsymbol{\alpha}_{i,11}$, and $\boldsymbol{\alpha}_{i,12}$ by c, c^2, and c^3, respectively. I am able to rule out this observational equivalence, because the NCES math exams in grades 10 and 12 are scored on the same vertical scale, which Agostinelli and Wiswall (2020) show implies that $\alpha_{10,j}^R = \alpha_{12,j}^R$ and $\mu_{10,j}^R = \mu_{12,j}^R$ for j equal to the NCES math exam for both URM and non-URM students.[24]

The mean of the latent factor is not separately identified from the mean of the measurements and is typically normalized to zero. Out of concern that a history of prior inputs produces a different initial distribution for different students, I let the mean of $\log K_{i,9}$ depend on a vector of initial conditions: $\mathbb{E}[\log K_{i,9} \mid \mathbf{W}_i] = \mathbf{W}_i'\mathbf{a}$. Note that it is not possible to identify a level shift in the mean of ninth grade measurements for under-represented minorities, $\boldsymbol{\mu}_9^{URM}$, from a shift in the initial mean of knowledge for under-represented minorities, $\mathbb{E}[\log K_{i,9} \mid URM_i = 1]$. A normalization is necessary, and I constrain the NCES math exams to have the same constants regardless of race: $\mu_{10,j}^{URM} = \mu_{10,j}^{WA}$ and $\mu_{12,j}^{URM} = \mu_{12,j}^{WA}$ for j equal to the NCES math exam. This way, $\mathbb{E}[\log K_{i,9} \mid URM_i = 1]$ can be identified separately from $\mathbb{E}[\log K_{i,9} \mid URM_i = 0]$.[25]

4.2 Identification of College Completion and Preference Parameters

College preferences for diversity, κ, are identified by the measurements of marginally admitted URM and white and Asian applicants. If, for example, the marginal URM admit to a specific school has lower GPAs and SAT scores than the marginal white and Asian admit, then this school must have a positive preference for diversity. This paper exploits variation in exam access and distance to college, which serve as exclusion restrictions that shift the probability of applying to college. Appendix C presents estimates from

[23]Theorem 5.1 in Anderson and Rubin (1956).

[24]I still allow for $\sigma_{10,j}^R$ to differ from $\sigma_{12,j}^R$ for the NCES math exams, so that the signal-to-noise ratios of the two exams may differ.

[25]Appendix B shows that the inferences drawn from the dynamic factor are robust to alternative normalizations.

first-stage regressions of applications to college on exam access (Table C-1) and distance to college (Table C-2). The regressions show that exam access increases applications to college, while distance to college affects the decision of which school to apply to. The exclusion restrictions provide reassurance that the model is not identified solely on the basis of functional form assumptions.

This paper will analyze how counterfactual admissions policies influence sorting to college and rates of college completion. The model generates estimates of treatment effects for college completion at school j relative to school k, $O_i(K_{i,12}, j) - O_i(K_{i,12}, k)$. These treatment effects are identified by randomness in admissions signals that causes students with the same $K_{i,12}$ to have different admissions sets and thus attend different colleges. Distributional plots of $\log K_{i,12}$ by school in section 6 reveal that there is considerable overlap in the knowledge distribution across colleges and thus sufficient support to analyze these treatment effects.

The scale of the extreme value shock for each school $\varepsilon_{i,c}$ is normalized to 1, and the mean value of not attending college has been normalized to zero for all students. However, the scale of the application cost shock, λ_A, has not been normalized. It is identified, because the benefit of the application portfolio in equation 13 is not multiplied by a parameter, and so the application choice does not face the usual observational equivalence in discrete choice models: It is not possible to increase the preference parameters and the scale parameter by a constant factor and leave the application probabilities in the likelihood function unchanged. In practice, it is important to let λ_A vary in optimization, particularly when the choice set is large, as is the case here. Constraining λ_A to be a small value, as is sometimes done in the literature on school choice, implies that the model can perfectly rationalize the large number of potential applications with a small number of parameters. Even after controlling for a rich set of observable characteristics of students and colleges, it is unlikely that any model could fully explain the heterogeneity in application patterns among American students.

5 Estimation

I estimate the model by deriving the likelihood function and using a Nested Fixed Point algorithm (NFXP). In estimation, I make use of the three NCES exams present in the ELS 2002, which are not observed by colleges, to aid in identification of the dynamic factor model. For clarity I distinguish between $\mathbf{y}_{i,10}$ and $\tilde{\mathbf{y}}_{i,10} = (\mathbf{y}_{i,10}, y_{i,10}^{(j)}, y_{i,10}^{(k)})$ for j and k equal to the NCES math and reading exams, and between $\mathbf{y}_{i,12}$ and $\tilde{\mathbf{y}}_{i,12} = (\mathbf{y}_{i,12}, y_{i,12}^{(j)})$ for j again equal to the NCES math exam. Colleges observe $\Omega_{i,12}$ while the econometrician observes

$$\tilde{\Omega}_{i,12} := \{\Omega_{i,9}, \tilde{\mathbf{y}}_{i,10}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,12}, h_{i,10}, h_{i,11}, h_{i,12}\} \ .$$

For each student, I observe the college attended, C_i, the admissions set, B_i, the application set A_i, their observed measurements, $(\mathbf{y}_{i,9}, \tilde{\mathbf{y}}_{i,10}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,12})$, their actions while in High School, a_i, and their initial conditions, $\Omega_{i,9}$. I also observe whether an individual graduates from college, $Complete_{i,c}$. Letting θ denote the entire set of model parameters, the likelihood contribution for individual i is

$$\begin{aligned} l_i(Complete_{i,c}, C_i, B_i, A_i, \tilde{\mathbf{y}}_{i,12}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,10}, \mathbf{y}_{i,9}, a_i \mid \Omega_{i,9}, \theta) =& P(Complete_{i,c} \mid C_i, \tilde{\Omega}_{i,12}, \theta) \times \\ & P(C_i \mid B_i, \Omega_{i,12}, \theta) \times \\ & P(B_i \mid A_i, \Omega_{i,12}, \theta) \times \\ & P(A_i \mid \Omega_{i,12}, a_i, \theta) \times \\ & f(\tilde{\mathbf{y}}_{i,12}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,10} \mid a_i, \Omega_{i,9}, \theta) \times \\ & P(a_i \mid \Omega_{i,9}, \theta) \times \\ & f(\Omega_{i,9}; \theta) \ , \end{aligned} \tag{20}$$

where $P(Complete_{i,c} \mid C_i, \tilde{\Omega}_{i,12}, \theta)$ comes from equation 19; $P(C_i \mid B_i, \Omega_{i,12}, \theta)$ comes from equation 10; $P(B_i \mid A_i, \Omega_{i,12}, \theta)$ comes from equation 11; $P(A_i \mid \Omega_{i,12}, a_i, \theta)$ comes from equation 15; $f(\tilde{\mathbf{y}}_{i,12}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,10} \mid \Omega_{i,9}, a_i, \theta)$ comes directly from the technology and measurement system in equations 1 and 2; $P(a_i \mid \Omega_{i,9}, \theta)$ is the solution to the problem of a high school student in 17; and $f(\Omega_{i,9}; \theta)$ are the initial conditions that vary with $\mathbf{W}_i$.

I choose θ to minimize the log-likelihood function:

$$L(\theta) = \sum_{i}^{N} \log l_i(Complete_{i,c}, C_i, B_i, A_i, \tilde{\mathbf{y}}_{i,12}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,10}, \mathbf{y}_{i,9}, a_i \mid \Omega_{i,9}, \theta) \ .$$

Several components of the log-likelihood function depend on different sets of parameters, so I first optimize partial likelihoods on the relevant set of parameters before optimizing the full likelihood. It is possible to obtain consistent estimates of the measurement and technology parameters by optimizing over the following partial likelihood,

$$\sum_{i=1}^{N} \log f(\mathbf{y}_{i,12}, \mathbf{y}_{i,11}, \mathbf{y}_{i,10} \mid a_i, \Omega_{i,9}, \theta) + \log f(\Omega_{i,9}; \theta) \ ,$$

which does not require solving for equilibrium in the college market. The second step is to search for the college completion parameters by maximizing

$$\sum_{i=1}^{N} \log f(O_i \mid C_i, \Omega_{i,12}, \theta) \ ,$$

which also does not require solving the college equilibrium. I then search over the preference parameters for both students and colleges by maximizing the remainder of the log likelihood function:

$$\sum_{i=1}^{N} \log P(C_i \mid B_i, \Omega_{i,12}, \theta) + \log P(B_i \mid A_i, \Omega_{i,12}, \theta) + \log P(A_i \mid \Omega_{i,12}, a_i, \theta) + \log P(a_i \mid \Omega_{i,9}, \theta)$$

After obtaining consistent estimates from maximizing the partial likelihoods, I optimize the full information likelihood to obtain a set of efficient estimates.

The College Market Equilibrium in section 3.6 assumes that colleges take as given the distribution of student test scores and applications, $\{\Omega_{i,12}, A_i\}_{i=1}^{N}$. Hence, despite the large number of application portfolios, solving for the equilibrium in the NFXP algorithm is not computationally costly: I need only to calculate the probabilities of admission and matriculation given the application/admission pairs observed in the data. However, computing $P(a_i \mid \Omega_{i,12}, \theta)$ *is* very costly, because for every student and every action they may take, I must integrate over the distribution of $(m_{i,12}, P_{i,12})$ given $(\Omega_{i,9}, a)$, and compute the inclusive value, $\overline{V}_i^{Coll}(m_{i,12}, P_{i,12}, SAT_i(a))$, which depends on all 584 possible application portfolios. To make the computation feasible, I precompute $\overline{V}_i^{Coll}(m_{i,12}, P_{i,12}, SAT_i)$ on a grid for each student and use interpolation to simulate the values between grid points. $\overline{V}_i^{Coll}(m_{i,12}, P_{i,12}, SAT_i)$ varies smoothly with $m_{i,12}$, and Monte Carlo simulations (available upon request) reveal that interpolation introduces negligible error. A similar interpolation is used to solve for the college market equilibrium in counterfactual simulations.

6 Results

6.1 Dynamic Factor Model

The parameters of the initial conditions distribution, in Table 4, reveal dramatic differences in initial knowledge across individuals. The mean and log variance of $\log K_{i,9}$ have been normalized to zero, so that the coefficients in the column labeled Mean can be interpreted in terms of standard deviations. The table shows that average knowledge in ninth grade is 0.65 sd lower for URMs relative to white and Asian students. I also find

Table 4: Parameters Governing Initial Distribution of Knowledge

	Mean	Log Variance
URM	-0.65	0.09
	(0.02)	(0.06)
Female	-0.02	-0.12
	(0.02)	(0.03)
Retain	-0.74	-0.25
	(0.04)	(0.08)
Single Parent	-0.13	0.01
	(0.02)	(0.05)
Mother: High School	0.24	-0.02
	(0.03)	(0.04)
Mother: Some College	0.41	-0.07
	(0.02)	(0.03)
Mother: Bachelors	0.71	0.02
	(0.03)	(0.05)
Mother: Postgraduate	0.90	0.05
	(0.05)	(0.05)
HH Income	0.08	-0.02
	(0.01)	(0.01)

The table presents estimates of parameters governing the initial distribution of skills in the ninth grade. The mean and variance have been normalized to 0 and 1, respectively, for individuals whose covariates are all equal to 0. Details regarding the distribution of knowledge are provided in section 3. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

that students who were retained prior to high school have three-quarters of a sd lower knowledge, students who grow up with a single mother lag behind by 0.13 sd, and initial knowledge is sharply increasing in mother's education. The coefficient on household income indicates that initial knowledge is higher by 0.08 sd for each additional \$100 spent on the child per week.[26]

The variance of the initial knowledge distribution differs by a subset of these variables. URMs have a higher variance of initial knowledge (by 0.09 log points) while variance is lower for girls and students growing up in richer households. Retained students have lower variance by 0.25 log points, consistent with them typically being selected from the left tail of the skill distribution.

[26] I assume, consistent with a range of estimates reviewed in Donni (2015), that families spend one quarter of their household income on the child. The median value of this variable in the data, 3, corresponds to a yearly income of \$62,400 ($3 \times 4 \times 100 \times 52 = 62400$).

Table 5: Estimates of $\boldsymbol{\mu}_t^R$

	URM	White & Asian	Difference
GPA, 9th grade	-0.20	-0.19	-0.01
	(0.03)	(0.02)	(0.02)
GPA, 10th grade	-0.21	-0.18	-0.03
	(0.02)	(0.02)	(0.02)
GPA, 11th grade	-0.23	-0.17	-0.06
	(0.02)	(0.02)	(0.02)
GPA, 12th grade	-0.28	-0.12	-0.16
	(0.02)	(0.02)	(0.02)
SAT Math	-0.69	-0.71	0.02
	(0.03)	(0.03)	(0.02)
SAT Verbal	-0.63	-0.63	0.00
	(0.03)	(0.03)	(0.02)
NCES Reading, 10th grade	-0.21	-0.22	0.01
	(0.03)	(0.02)	(0.02)
NCES Math, 10th grade	-0.25	-0.25	0
	(0.03)	(0.03)	(-)
NCES Math, 12th grade	-0.25	-0.25	0
	(0.03)	(0.03)	(-)

The table displays estimates of $\boldsymbol{\mu}_t^R$ in equation 2. Details regarding the the measurement system are provided in section 3. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

Table 5 displays estimates of $\boldsymbol{\mu}_t^R$ for $R = URM, WA$. The table can be used to assess whether grades and exams are biased against URMs, as $\boldsymbol{\mu}_t^R$ governs level shifts in the measurements across demographic groups after controlling for knowledge. Recall that the identifying normalization discussed in section 4.1, $\mu_{10,j}^{URM} = \mu_{10,j}^{WA}$ for j equal to the NCES math exam, rules out bias in this exam. The numbers in Table 5 should therefore be interpreted as bias relative to this exam.[27] There does not appear to be evidence that the SAT is biased against URMs. In fact, URMs score marginally higher on all of the standardized tests than might be expected conditional on knowledge. The estimated parameters in Table 5 suggest that, if anything, eleventh and twelfth grade GPAs are more biased against URM students than are standardized exams.[28]

[27] Appendix B provides additional analysis to show that the inferences in the table are robust to alternative normalizations.

[28] Implicit bias among teachers, as measured by the Implicit Association Test, has been shown by Carlana

Table 6: Signal-to-Noise Ratios ($\frac{\alpha_{t,j}}{\sigma_{t,j}}$)

	URM	White & Asian	Difference
GPA, 9th grade	0.67	0.93	-0.26
	(0.02)	(0.01)	(0.04)
GPA, 10th grade	0.68	0.92	-0.25
	(0.02)	(0.01)	(0.03)
GPA, 11th grade	0.60	0.80	-0.20
	(0.01)	(0.01)	(0.09)
GPA, 12th grade	0.48	0.60	-0.12
	(0.02)	(0.01)	(0.06)
SAT Math	1.65	1.87	-0.22
	(0.03)	(0.02)	(0.02)
SAT Verbal	1.29	1.25	0.04
	(0.06)	(0.01)	(0.04)
NCES Reading, 10th grade	1.13	1.04	0.08
	(0.02)	(0.01)	(0.03)
NCES Math, 10th grade	1.83	1.65	0.17
	(0.03)	(0.02)	(0.04)
NCES Math, 12th grade	2.06	1.92	0.14
	(0.04)	(0.01)	(0.05)

The table displays estimates of signal-to-noise ratios, $\frac{\alpha_{t,j}}{\sigma_{t,j}}$, for all measurements in the data. The NCES exams are used to identify the technology of skill formation, but are not available to colleges when determining whom to admit. More details regarding the the measurement system are provided in section 3. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

Even if the SAT is not biased against URMs, its informativeness as a signal may still vary across demographic groups. Table 6 presents estimates of signal-to-noise ratios for all the measurements in the data. For a given measurement, $y_{t,j}$, the signal-to-noise ratio is computed as $\alpha_{t,j}/\sigma_{t,j}$. A signal-to-noise ratio that exceeds unity indicates that the measurement contains more information than noise. The table indicates that GPAs become worse signals in later years of high school. It also shows that the standardized exams convey significantly greater information than GPAs, with the math portion of the SAT and the math exams administered by the NCES being particularly informative. GPAs in every grade are less informative for URMs than for white and Asian students. The math

(2019) and Van den Bergh et al. (2010) to predict both gender and racial test scores gaps and could be a source of the GPA biases seen here.

Table 7: Technology of Skill Formation

	URM	White & Asian
Knowledge(-1)	0.99	1.02
	(0.01)	(0.02)
Study, 10 hours/wk	0.10	0.08
	(0.01)	(0.02)
Private School	0.05	0.01
	(0.01)	(0.02)
Free Lunch	-0.14	-0.20
	(0.04)	(0.04)
Student Teacher Ratio	0.07	-0.07
	(0.08)	(0.17)
Mother: High School	-0.02	-0.02
	(0.02)	(0.03)
Mother: Some College	-0.03	-0.03
	(0.02)	(0.02)
Mother: Bachelors	-0.04	-0.04
	(0.02)	(0.03)
Mother: Postgraduate	-0.04	-0.04
	(0.02)	(0.05)

The table displays estimates of parameters governing the technology of skill formation. Study refers to the effect of studying 10 hours per week on next year's skills. More details on the technology are provided in section 3. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

portion of the SAT is also a worse signal for URMs, while the other standardized tests are marginally more informative for URMs.

The parameters of the measurement system therefore indicate that the SAT does not appear to be biased against URMs. But, there is merit to the concern that the math portion of the SAT may not be as informative for a URM student as it is for a white or Asian student. The same can also be said for grades in school. Much of the literature on grading practices, for example Botelho, Madeira, and Rangel 2015 and Rauschenberg 2014, has focused on the first moments of grades. The results presented here suggest that second moments may also vary across demographic groups.

Table 7 presents estimates of the technology of skill formation. I find no evidence of differences in this technology for URMs and white and Asian students. Both have an autoregressive parameter for knowledge equal to one, indicating that knowledge does not depreciate throughout high school. This suggests that it may be difficult for students who enter high school with a low level of knowledge to catch up to their peers by the time they apply to college. I also find that an additional ten hours of study time per week increases knowledge by 0.08–0.10 sd. Hence, a URM student who studies 10 hours per week will, all else equal, improve her skills by 0.30 sd between the end of 9th grade and the end of high school. Studying may therefore deliver significant marginal returns for students whose initial conditions place them on the cusp of gaining admission to college. Students at poorer high schools, as indicated by the proportion of students qualifying for a free or reduced-price lunch, accumulate less knowledge. The effect of class size on skill development is insignificant. Attending a private school, either catholic or nondenominational, has a small positive effect on knowledge accumulation for URMs. In contrast to the effects of mother's education on initial knowledge, students with more highly educated mothers tend to regress slightly toward the mean throughout high school.

6.2 Estimated Preferences

Table 8 presents estimates of the preference parameters for students and colleges. The fixed effects for each college are ordered in a way that is consistent with selective universities being more highly valued. Private colleges (tiers 1-3) are also more preferred than public colleges (tiers 4-6) with a similar degree of selectivity. Distaste for distance is similar for both private and public colleges and equal to three-quarters of the magnitude of the coefficient on tuition, indicating that students are indifferent between attending a college that is 100 miles closer and paying tuition that is higher by \$75 per week (\$3,900 per year). Students are more likely to attend a school where a higher proportion of the cost of attendance is covered by financial aid. They also have a strong preference for college completion. Comparison with the fixed effects, $\overline{U}_1, \ldots, \overline{U}_6$, reveals that college is highly valued because of the degree that it confers rather than because of amenities unrelated to the degree. A comparison between the coefficients on tuition and completion suggests that a college degree is valued at $5.12/0.16 \times 100 = \$3,200$ dollars per week, or $\$166,400$ per year.

Students dislike studying, but studying is less costly if they come from richer households.[29] Likewise, SAT-taking is costly but greater income and logistical access make it less so. To put some of the numbers in Table 8 in perspective, the disutility of studying ten hours a week for a student from a median income family is $(-0.12 + 0.01 \times 3.00) \times$

[29]This is consistent with poorer students having a higher opportunity cost for studying. The probability of working for pay while in high school is negatively associated with family income in the ELS 2002.

Table 8: Preference Parameters

	Value	Standard Error
Student Preferences for Universities		
College 1, $\overline{U}_1$	-0.04	0.19
College 2, $\overline{U}_2$	-1.62	0.18
College 3, $\overline{U}_3$	-1.83	0.18
College 4, $\overline{U}_4$	-0.96	0.27
College 5, $\overline{U}_5$	-2.16	0.15
College 6, $\overline{U}_6$	-2.29	0.12
Portfolio Shock, Scale	1.14	0.03
Distance, Public	-0.11	0.08
Distance, Private	-0.13	0.08
Tuition	-0.16	0.03
Aid	0.59	0.10
Completion	5.12	0.10
Student Preferences in High School		
Hours, γ_H	-0.12	0.00
Hours $\times$ Income, γ_H^{Inc}	0.01	0.00
SAT, γ_S	-1.48	0.07
SAT $\times$ Access, γ_S^Z	0.09	0.02
SAT $\times$ Income, γ_S^{Inc}	0.17	0.02
College Preferences		
Preference for Knowledge, Private	0.89	0.00
Preference for Diversity, Private	0.11	
Preference for Knowledge, Public	0.91	0.00
Preference for Diversity, Public	0.09	

The table displays estimates of preference parameters of students and colleges in the model. More details are provided in section 3. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

$10 = -0.90$. This equates to $-0.90/0.16 \times 100 = 562.50$ dollars a week, and is worth $0.90/5.12 = 0.18$ of a college degree. This median student would therefore study 10 additional hours a week if doing so increased his probability of attaining a college degree by 18% or more.

Private colleges value knowledge less and diversity more than public colleges: Eleven-hundredths of private universities' utility is derived from the diversity of its student

Table 9: Admissions Thresholds

	Raw		*Standardized*	
	White/Asian	URM	White/Asian	URM
Tier 1	1.59	0.92	1.00	0.42
Tier 2	0.56	-0.06	0.10	-0.43
Tier 3	-0.01	-0.57	-0.39	-0.88
Tier 4	1.11	0.61	0.58	0.15
Tier 5	0.56	0.09	0.10	-0.31
Tier 6	0.07	-0.38	-0.33	-0.72

The table displays estimated admissions thresholds for each tier of college. Students who apply to a school with an application signal that exceeds their demographic-specific threshold are granted admission. Columns labeled Raw indicate the threshold in terms of $\log K_{i,12}$, while in the columns labeled Standardized, the thresholds have been normalized by subtracting the mean dividing by the standard deviation of $\log K_{i,12}$. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

body, while the corresponding fraction for public universities is nine out of one hundred. This difference is statistically significant. The weight placed on diversity causes a wedge to arise between the two admissions thresholds at each college. Table 9 displays these thresholds scaled by the population mean and standard deviation of $\log K_{i,12}$. The table shows that colleges that are more highly valued according to the fixed effects in Table 8 have higher admissions thresholds. Additionally, for each college tier, the threshold is lower for URM than for white and Asian students.

Table 10 shows estimates of the application cost parameters. Fixed costs – which vary by income, mother's education, distance, and a tier-specific constant – capture both monetary and nonmonetary deterrents to applications. A single application to a tier one school is over five times as costly as an application to a tier six school, all else equal. The signs on income are consistent with Hoxby and Avery (2012) who find that higher income students prefer to apply to state flagships and dislike nonselective schools. At the same time, the evidence in Table 10 suggests that nonmonetary costs are more salient. Having a mother with a college degree dramatically reduces the fixed cost of applying to a tier one school. This effect, -1.05, is greater than the effect of moving from the poorest to the richest household in the sample $(-0.1 \times (9.62 - 0.24)) = -0.938$. Distance is also an important determinant of where students apply, especially at lower-ranked universities.

Marginal costs, which are modeled as a fraction of fixed costs, are lowest at the most selective schools (tiers one and four). These schools have high fixed costs, indicating

Table 10: Application Cost Parameters

	Fixed Cost				*Marginal Cost*
	Constant	Income	Mother's Ed	Distance	Fraction
Tier 1	4.26	-0.10	-1.12	0.29	0.02
	(0.17)	(0.02)	(0.11)	(0.05)	(0.02)
Tier 2	2.82	-0.10	-0.74	0.49	0.40
	(0.10)	(0.01)	(0.06)	(0.07)	(0.03)
Tier 3	2.16	0.02	-0.19	1.44	0.65
	(0.08)	(0.01)	(0.06)	(0.13)	(0.03)
Tier 4	2.83	-0.09	-0.73	0.52	0.38
	(0.11)	(0.01)	(0.07)	(0.05)	(0.03)
Tier 5	1.78	-0.07	-0.37	0.55	1.11
	(0.07)	(0.01)	(0.05)	(0.05)	(0.05)
Tier 6	0.77	0.02	0.07	1.22	0.71
	(0.05)	(0.01)	(0.04)	(0.09)	(0.05)

The table displays estimates of parameters governing the cost of applying to college. Mother's Ed is an indicator for whether an individual's mother has a bachelor's degree. Marginal costs are a fraction of fixed costs. Fixed and marginal costs are school-specific. More details on application costs are provided in section 3. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

that it is relatively rare for a student to apply to them, but conditional on sending at least one application, the low marginal cost indicates that many students send multiple applications to schools within these tiers.

6.3 College Completion

Table 11 displays estimates of the parameters governing college completion. The estimates reveal that greater knowledge when enrolling in college increases the probability of completion at all schools. Tier six schools are the only ones with a statistically significant constant, indicating that completion is lower at these schools after controlling for knowledge. Although URMs are 11 pp less likely to complete college conditional on enrolling (Table 2), I find that there is no significant difference in completion rates by demographic group after controlling for knowledge and the college attended. I also find that having a college-educated mother has no effect on college completion and that family income is marginally significant. The results suggest that, if colleges could enroll URMs with similar skill levels as white and Asian matriculants, the college completion gap would disappear.

Table 11: College Completion Model

	Estimate	Standard Error
Tier 1	-0.03	0.25
Tier 2	0.22	0.22
Tier 3	0.01	0.15
Tier 4	0.23	0.26
Tier 5	-0.02	0.16
Tier 6	-0.29	0.08
$\log K_{i,12} \times$ Tier 1	0.28	0.10
$\log K_{i,12} \times$ Tier 2	0.28	0.04
$\log K_{i,12} \times$ Tier 3	0.31	0.03
$\log K_{i,12} \times$ Tier 4	0.20	0.05
$\log K_{i,12} \times$ Tier 5	0.35	0.04
$\log K_{i,12} \times$ Tier 6	0.43	0.02
URM	-0.02	0.06
Income	0.02	0.01
Mother has college degree	0.06	0.06

The table displays parameters of the college completion model, equation 18. The constant and slope with respect to knowledge, $\log K_{i,12}$, are allowed to vary by college tier. More details on college completion are provided in section 3. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

6.4 Goodness of Fit

Table 12 shows that the model successfully replicates many data moments. The model does a good job of matching the proportions of students at each type of college, as well as the overall percentage of students, 43%, attending any four-year university. The model somewhat overpredicts attendance by URMs at tiers 2 and 5 and underpredicts it at tiers 3 and 6. The model also closely replicates rates of college completion conditional on enrollment, both overall (68% in the data, 69% in the model) and by demographic groups (59% for URMs and 70% for whites and Asians in the data versus 62% for URMs and 71% for whites and Asians in the model).

The model is able to replicate the clear pattern of sorting by SAT scores to universities that exists in the data. The third panel of Table 12 shows mean standardized scores on both the math and verbal SAT examinations for students who attend each tier of college.

The model closely replicates hours spent studying across all students in the sample: 6.21 hours in the data versus 6.24 hours in model simulations. However, the model somewhat overpredicts study time by URMs (5.69 hours in the data versus 6.00 hours in the model) and underpredicts it among white and Asian students (6.41 hours in the data versus 6.33 hours in the model). A similar story holds for SAT-takeup rates: The model replicates take-up overall, but slightly overpredicts take-up for URMs (by 2 pp). Put another way, by controlling for family income, exam access, and the chance of admission to college, this model can explain 80% of the gap in SAT take-up by race.

The model matches overall patterns of sorting to college by family income. Family income affects both the initial conditions of the skill distribution as well as the cost of applying to college, so the model can quite precisely target the skill and income of applicants to each tier of college. What the model is unable to replicate is the finding that URM students at elite private universities (tier 1) come from lower income families than those attending tier 2 universities. This feature of the data may be due to aggressive outreach by elite colleges to magnet schools in urban areas as documented in Hoxby and Avery (2012). This sort of targeting is outside the scope of the model.

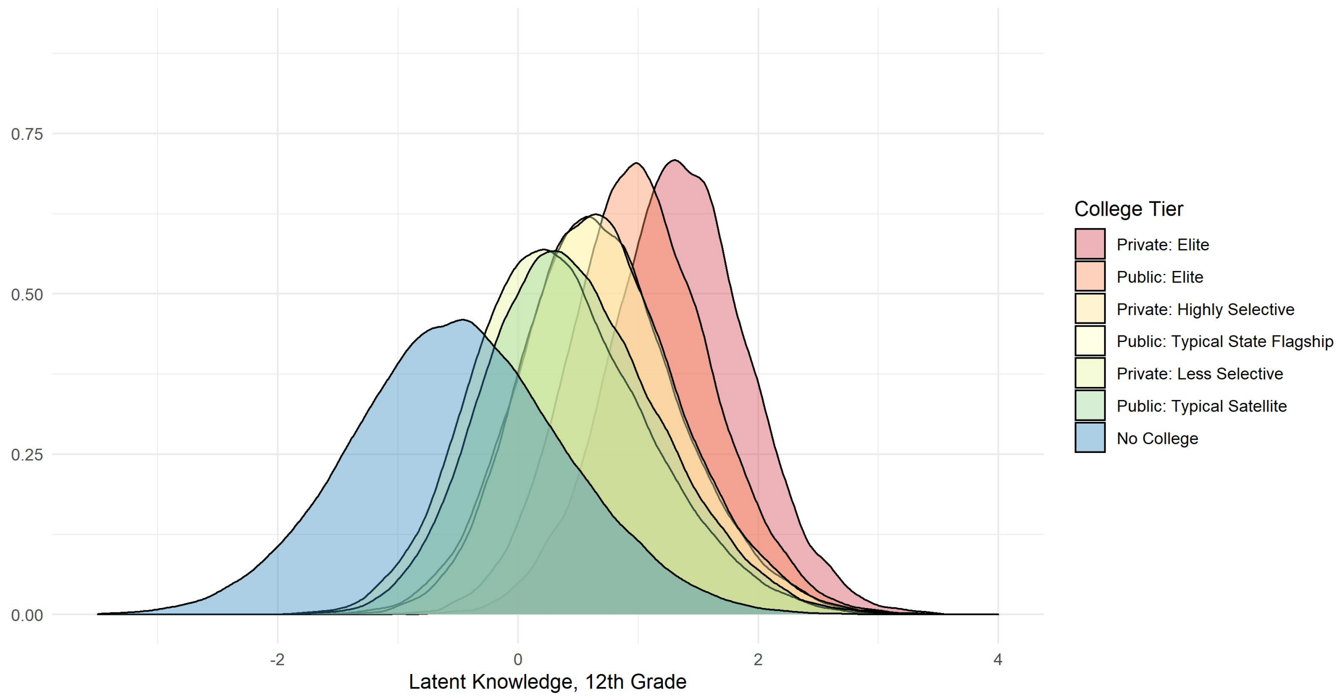
Figure 4, which plots densities of $\log K_{i,12}$ by the school attended, depicts patterns of sorting by knowledge across college tiers. The figure reinforces the patterns of sorting by SAT scores seen in Table 12, as there is a definite ordering to the peaks of each density. The figure also reveals substantial overlap in the knowledge distribution at all colleges, even between students who attend no college and those who attend elite colleges. In the next section, I will evaluate counterfactual policies that may send a new pool of students to college. The overlap of ability across colleges provides reassurance that the predictions of college completion in the next section are supported by the data .

Table 12: Goodness of Fit

	Data			Model		
	All	URM	WA	All	URM	WA
Fraction Attending						
Tier 1	0.019	0.016	0.020	0.025	0.017	0.028
Tier 2	0.072	0.035	0.086	0.077	0.057	0.085
Tier 3	0.066	0.068	0.065	0.06	0.047	0.065
Tier 4	0.048	0.032	0.054	0.054	0.038	0.060
Tier 5	0.063	0.022	0.079	0.066	0.043	0.075
Tier 6	0.165	0.160	0.167	0.149	0.108	0.165
Any College	0.433	0.333	0.47	0.432	0.310	0.478
SAT Math Score						
Tier 1	1.20	0.68	1.35	1.04	0.49	1.16
Tier 2	0.56	0.48	0.57	0.34	-0.16	0.46
Tier 3	-0.10	-0.72	0.14	0.01	-0.43	0.14
Tier 4	0.80	0.40	0.89	0.70	0.25	0.80
Tier 5	0.38	-0.16	0.43	0.37	-0.07	0.46
Tier 6	-0.04	-0.42	0.10	0.13	-0.3	0.24
No College	-0.58	-0.92	-0.45	-0.88	-1.21	-0.74
SAT Verbal Score						
Tier 1	1.25	0.88	1.37	0.96	0.50	1.06
Tier 2	0.69	0.58	0.70	0.32	-0.12	0.43
Tier 3	-0.04	-0.48	0.14	0.03	-0.37	0.14
Tier 4	0.76	0.43	0.84	0.65	0.27	0.74
Tier 5	0.34	-0.18	0.40	0.35	-0.04	0.43
Tier 6	-0.03	-0.39	0.10	0.14	-0.24	0.24
No College	-0.50	-0.84	-0.38	-0.79	-1.11	-0.64
Household Income						
Tier 1	103K	69K	113K	105K	92K	108K
Tier 2	84K	78K	85K	90K	77K	93K
Tier 3	73K	59K	79K	73K	59K	77K
Tier 4	102K	91K	105K	99K	87K	101K
Tier 5	86K	61K	89K	89K	74K	92K
Tier 6	69K	56K	74K	72K	60K	75K
No College	54K	43K	60K	52K	40K	58K
Complete College	0.678	0.594	0.701	0.689	0.619	0.706
Hours Study	6.21	5.69	6.41	6.24	6.00	6.33
Take SAT	0.76	0.66	0.79	0.76	0.68	0.79

The table compares moments in the data with their model counterparts by simulating the model according to the estimated parameters. Simulated moments are computed using 200 simulated data sets. WA refers to the population of white and Asian students. SAT scores are normalized by the population mean and standard deviation in the data. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

Figure 4: Latent Knowledge, by College Attended



The figure shows the simulated distribution of $\log K_{i,12}$ by college attended. Densities are computed using 200 simulated data sets.

7 Counterfactuals

In this section, I evaluate several counterfactual policies. The first is a policy that bans the SAT. In this counterfactual, colleges rely on grades and the variables governing the initial skill distribution when determining whom to admit. Students can respond to the policy by changing where they apply to college and by studying more or less while in high school. The choice of whether to take the SAT is eliminated from students' choice set. Universities optimally set admissions thresholds so that the expected demand for university enrollment equals capacity. A second policy mandates all students take the SAT. As before, students respond along the application and study margins, and colleges respond by adjusting thresholds.

7.1 Main Findings

Table 13 shows that eliminating the SAT slightly reduces URM enrollment at elite public and private colleges, while mandating it slightly increases URMs overall. Both policies generate nonnegative changes in the enrollment of students from below median-income

Table 13: Access to College

	URM Attendance			Low-Income Attendance		
	Status Quo	No SAT	SAT-for-All	Status Quo	No SAT	SAT-for-All
Tier 1	0.017	0.016	0.017	0.010	0.010	0.010
Tier 2	0.057	0.057	0.057	0.044	0.045	0.046
Tier 3	0.047	0.048	0.051	0.049	0.051	0.052
Tier 4	0.038	0.036	0.038	0.024	0.026	0.026
Tier 5	0.043	0.041	0.043	0.04	0.042	0.043
Tier 6	0.108	0.107	0.113	0.123	0.128	0.131
Any College	0.310	0.305	0.319	0.289	0.302	0.309

The table displays the rates of attendance for URM and low-income students at each college tier under three separate policies: the status quo, a policy where the SAT is banned, and a policy in which all students take the SAT and submit the scores with their applications. Low-income refers to students whose families earn less than the median ($52,500 per year). Simulated moments are computed using 200 simulated data sets. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

families at all colleges.[30] The reason why banning the SAT fails to raise URM enrollment is due to the offsetting effects of two separate phenomena, as shown in Table 14. Removal of the SAT barrier causes a rise in applications to college. This increase is larger for URMs, who are less likely to take the exam, than for white and Asian students (30% versus 17%). White and Asian students, however, send more applications at baseline (1.60 per student versus 1.20), and so the fraction of total college applications coming from URMs rises only slightly, from 22% to 24%. This might be large enough to raise URM enrollment were it not for another phenomenon. URMs who are induced to apply to college after banning the SAT are weaker candidates than marginal white and Asian applicants. The value of $\log K_{i,12}$ for the average URM applicant declines by 0.16 sd, while it declines by 0.12 sd for white and Asian applicants. The combination of more URM applicants but lower average quality of their applications leads to no change in URM enrollment.

The No-SAT and SAT-for-All policies have little effect on URM enrollment, but they cause changes along other dimensions. Figure 5 shows that the two policies have considerably different effects on college sorting by knowledge. Eliminating the SAT renders the distributions of knowledge more homogeneous across colleges, while mandating the SAT has the opposite effect. By giving each college access to two additional high-quality signals in the SAT math and verbal exams, colleges are better able to draw an inference on each applicant's latent knowledge. This allows more sought-after schools to better select

[30]Median family income in the estimation sample is $52,500.

Table 14: Application Patterns

	Status Quo			No SAT			SAT-for-All		
	All	URM	WA	All	URM	WA	All	URM	WA
Apps per student	1.49	1.20	1.60	1.79	1.57	1.88	1.77	1.57	1.85
Application Growth (%)				20	30	17	19	31	15
% Applications from URMs		22			24			24	
$\mathbb{E}[\log K_{i,12}]$ per applicant	0.74	0.12	0.93	0.60	-0.04	0.81	0.61	-0.01	0.83

The table displays application statistics by demographic group under three separate policies: the status quo, a policy that bans the SAT, and a policy in which all students take the SAT and submit their scores with their applications. Simulated moments are computed using 200 simulated data sets. WA refers to the population of white and Asian students. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

highly-skilled applicants for admission, leading to greater assortative matching.

Changes in assortative matching cause changes in college completion. Table 15 shows that banning the SAT causes completion rates at tiers 1 and 4 to fall by as much as 3 pp. Less selective schools instead experience higher graduation rates, as they enroll stronger students who are turned away by elite schools because of noisier application signals. By contrast, the SAT-for-All policy increases completion at lower ranked schools without lowering it at elite schools. Rather than drawing students away from elite colleges, the SAT-for-All policy enables schools in tiers 2, 3, 5, and 6 to identify qualified students for admission among those who did not take the SAT and thus did not apply to college in the status quo. College completion rises by 1.4–1.6 pp overall.

The results show that elite private and public colleges (tiers 1 and 4) have the most to lose from a policy that bans the SAT. Average knowledge, college completion, and URM attendance all decline at these schools. The increase in signal variance causes all colleges to inadvertently reject skilled candidates, however less selective colleges are able to enroll students rejected by higher tiers, but there is no higher tier for elite colleges to draw from. Hence, they experience the largest declines in the skill of their student body.

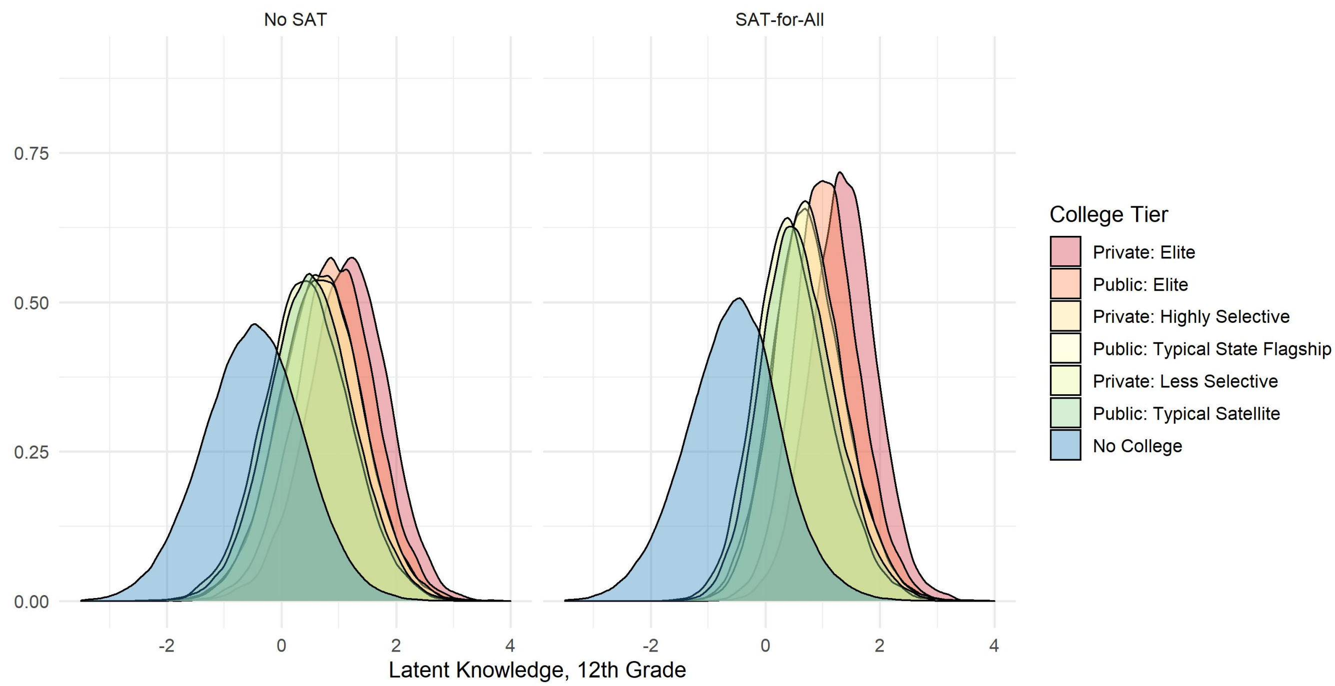
The mean of knowledge across all four-year colleges does not change under the No-SAT policy, while the SAT-for-All policy results in an increase in $\log K_{i,12}$ of 0.12 (0.09) sd for URM (white and Asian) matriculants. It seems surprising that removing the SAT does not lower the mean knowledge of college students. One might expect that the lower signal quality provided by grades would lead to the admission of weaker students. The next section explains that the equilibrium response by colleges forestalls this outcome.

Table 15: Counterfactuals

	Status Quo			No SAT			SAT-for-All		
	All	URM	WA	All	URM	WA	All	URM	WA
Complete College									
Tier 1	0.761	0.694	0.776	0.734	0.669	0.748	0.759	0.701	0.773
Tier 2	0.745	0.687	0.760	0.750	0.686	0.766	0.761	0.698	0.777
Tier 3	0.645	0.57	0.665	0.648	0.571	0.670	0.655	0.584	0.677
Tier 4	0.757	0.718	0.766	0.746	0.712	0.754	0.763	0.722	0.772
Tier 5	0.722	0.666	0.734	0.721	0.652	0.735	0.730	0.676	0.741
Tier 6	0.626	0.538	0.648	0.64	0.555	0.661	0.651	0.571	0.671
All Schools	0.689	0.619	0.706	0.692	0.619	0.709	0.703	0.635	0.72
Household Income									
Tier 1	105K	92K	108K	103K	91K	106K	103K	90K	106K
Tier 2	90K	77K	93K	88K	77K	91K	87K	73K	90K
Tier 3	73K	59K	77K	71K	59K	74K	71K	58K	75K
Tier 4	99K	87K	101K	97K	87K	99K	96K	85K	99K
Tier 5	89K	74K	92K	87K	73K	90K	86K	70K	89K
Tier 6	72K	60K	75K	71K	60K	73K	70K	58K	73K
No College	52K	40K	58K	53K	41K	59K	54K	41K	60K
Attend Any College	0.432	0.310	0.478	0.432	0.305	0.481	0.432	0.319	0.476
Hours Study	6.24	6.00	6.33	6.36	6.14	6.45	6.35	6.11	6.45
$\mathbb{E}[\log K_{i,12}]$: Any 4-yr College	0.65	0.20	0.76	0.65	0.18	0.76	0.74	0.32	0.85
$\mathbb{E}[\log K_{i,12}]$: No College	-0.49	-0.86	-0.31	-0.49	-0.84	-0.32	-0.56	-0.92	-0.39

The table displays summary statistics under three separate policies: the status quo, a policy that bans the SAT, and a policy in which all students take the SAT and submit their scores with their applications. Simulated moments are computed using 200 simulated data sets. WA refers to the population of white and Asian students. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

Figure 5: Latent Knowledge Distribution, Counterfactuals



The figure shows the simulated distributions of $\log K_{i,12}$ by college attended under two counterfactual policies. The first eliminates the SAT in college admissions, while the second mandates that every high school student take the SAT and submit their scores with their college application. Densities are computed using 200 simulated data sets.

7.2 Model Mechanisms

Eliminating the SAT causes four changes in the model: a shift in the set of measurements used to determine admission, endogenous applications, endogenous study decisions by high school students, and reoptimization by capacity-constrained colleges. To understand the quantitative importance of each of these elements, I simulate the model five times, starting from the status quo and adding one element at a time until I arrive at the full No-SAT counterfactual. Summary statistics for five major variables – URM attendance, knowledge, household income, college completion, and total attendance – under each model simulation are presented in Table 16. The second column, labeled No Sat, holds applications and study effort fixed and removes the SAT from among the set of measurements used to determine admission. The third column allows for endogenous applications, while the fourth column allows for endogenous applications together with endogenous effort. The final column imposes equilibrium in the college market.

A clear pattern emerges from the analysis. Knowledge and household income become more equalized across colleges in the No SAT and No SAT + Endogenous Applications simulations. The spread of $\log K_{i,12}$ between students at top tier colleges and those not attending college falls by nearly 0.5 sd. The percentage of URMs attending college rises dramatically, from 31.0% to 39.2%. However, the introduction of endogenous effort raises sorting by knowledge and household income. The imposition of equilibrium furthers this stratification, and dramatically reduces access to college by URMs, as marginal entrants are shut out of college by higher admissions standards. The last line in the table shows how, in the absence of the capacity constraints imposed in equilibrium, college attendance would be 8.2 pp higher.

The reason why endogenous effort increases assortative matching in partial equilibrium is subtle. When the SAT is eliminated, average effort increases from 6.21 to 6.34 hours a week. This increase in average effort masks two countervailing changes: Students who did not take the SAT in the status quo increase their study hours dramatically, from 4.25 to 5.93 hours. But, the reduction in signal quality arising from the SAT's elimination causes those formerly taking the SAT to reduce their hours worked, from 6.85 to 6.47 hours. Reduction in study hours pulls in the right tail of $\log K_{i,12}$, which explains why enrollment in Tier 1 colleges falls between columns 3 and 4 of Table 16. Students from families with high incomes reduce study hours by less, because effort is less costly for them, which results in a modest strengthening of the correlation between household income and $\log K_{i,12}$, from 0.35 to 0.36, and a rise in assortative matching.

When college sorting by knowledge changes, so do rates of college completion. Banning the SAT and allowing for endogenous applications reduce completion at all colleges

Table 16: Model Mechanisms

	Status Quo	No Sat	+ Endogenous Apps	+ Endogenous Effort	+ Equilibrium
URM Attendance					
Tier 1	0.017	0.015	0.023	0.018	0.016
Tier 2	0.057	0.037	0.068	0.071	0.057
Tier 3	0.047	0.067	0.058	0.063	0.048
Tier 4	0.038	0.036	0.047	0.043	0.036
Tier 5	0.043	0.033	0.053	0.055	0.041
Tier 6	0.108	0.128	0.142	0.148	0.107
Any College	0.310	0.317	0.392	0.399	0.305
$\mathbb{E}[\log K_{i,12}]$					
Tier 1	1.35	1.03	0.88	1.14	1.15
Tier 2	0.68	0.66	0.50	0.57	0.67
Tier 3	0.36	0.23	0.26	0.27	0.44
Tier 4	1.03	0.83	0.72	0.87	0.92
Tier 5	0.71	0.61	0.52	0.59	0.68
Tier 6	0.47	0.34	0.32	0.36	0.53
No College	-0.49	-0.39	-0.47	-0.55	-0.49
HH Income					
Tier 1	105K	100K	94K	102K	103K
Tier 2	90K	85K	83K	85K	88K
Tier 3	73K	73K	67K	67K	71K
Tier 4	99K	97K	90K	96K	97K
Tier 5	89K	84K	82K	84K	87K
Tier 6	72K	70K	66K	67K	71K
No College	52K	54K	54K	52K	53K
Complete College					
Tier 1	0.761	0.716	0.700	0.731	0.734
Tier 2	0.745	0.747	0.726	0.735	0.750
Tier 3	0.645	0.618	0.619	0.630	0.648
Tier 4	0.757	0.738	0.730	0.740	0.746
Tier 5	0.722	0.711	0.700	0.705	0.721
Tier 6	0.626	0.607	0.600	0.607	0.640
All Colleges	0.689	0.671	0.663	0.669	0.692
Total Attendance					
Tier 1	0.025	0.027	0.032	0.025	0.025
Tier 2	0.077	0.070	0.089	0.091	0.077
Tier 3	0.060	0.057	0.068	0.074	0.060
Tier 4	0.054	0.053	0.065	0.060	0.054
Tier	0.066	0.064	0.076	0.078	0.066
Tier 6	0.149	0.153	0.177	0.186	0.149
Any College	0.432	0.424	0.507	0.514	0.432

Each column in the table presents moments from a different simulation. The leftmost column simulates the status quo policy. The second column removes the SAT from admissions but holds applications and study effort fixed. The third column lets applications respond endogenously. The fourth column lets study effort respond endogenously, and the final column imposes equilibrium in the college market. Simulated moments are computed using 200 simulated data sets. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

by up to 6.1 pp. The introduction of endogenous effort and college optimization in equilibrium instead raises completion. In the full equilibrium without the SAT, elite private colleges have lower rates of completion, because noisier application signals cause them to lose out on some highly-skilled candidates, who then enroll in less selective universities. This, together with the admission of some strong students who do not take the SAT in the status quo, causes graduation rates at public satellite colleges (tier 6) to rise.

The trend moving from left to right in Table 16 of decreasing and then increasing stratification suggests that a lack of income at the application margin does not pose a prohibitive barrier to college access. The model allows the cost of application to vary with family income, but allowing for endogenous applications actually reduces stratification by income. Instead, unequal pre-college human capital investment generates a distribution of cognitive skills that results in children attending markedly different colleges based on their income. Restrictive supply at four-year colleges exacerbates this trend.

7.3 SAT-Optional Policy at Elite Colleges Only

Because I find that elite colleges suffer the most from banning the SAT, in this section I evaluate a policy in which elite private and public colleges allow students to apply whether they have an SAT score or not, while other colleges continue to require the SAT. This policy gives students who do not take the exam the option to apply to either elite colleges or to no college at all. For simplicity, I abstract away from the strategic decision of whether an SAT-taker should send an SAT score and instead assume that all students who take the SAT send their scores.

Table 17 shows the pattern of college sorting by knowledge under the SAT-Optional policy. The first two sets of columns show sorting under the status quo and under the No-SAT policy analyzed previously for comparison. Relative to the status quo, there is little change in sorting by knowledge to college when elite colleges go SAT-Optional. Average knowledge at tier one schools falls only marginally, while it actually increases from 1.03 to 1.10 sd above the mean at elite public colleges (tier four). This result contrasts with the decline in average knowledge under the No-SAT policy, and it arises because elite colleges receive applications from students who applied in the status quo with an SAT score, for whom there is no loss of information, plus additional applications from students who did not take the SAT. Because some of these new applicants have quite high admissions signals, admissions thresholds at tier one (tier four) schools increase by nearly one-tenth (two-tenths) of a standard deviation (Table 18). The effect of raising thresholds offsets the cost of not having the SAT for some of the applicants and leaves elite colleges with

Table 17: SAT-Optional at Elite Schools

	Status Quo			No SAT			SAT-Optional		
$\mathbb{E}[\log K_{i,12}]$	All	URM	WA	All	URM	WA	All	URM	WA
Tier 1	1.35	0.87	1.46	1.15	0.63	1.27	1.34	0.82	1.46
Tier 2	0.68	0.21	0.80	0.67	0.20	0.79	0.67	0.21	0.78
Tier 3	0.36	-0.09	0.48	0.44	-0.03	0.58	0.34	-0.11	0.46
Tier 4	1.03	0.63	1.12	0.92	0.47	1.01	1.10	0.65	1.20
Tier 5	0.71	0.29	0.80	0.68	0.23	0.78	0.69	0.26	0.79
Tier 6	0.47	0.04	0.57	0.53	0.08	0.63	0.44	0.02	0.55
Any College	0.65	0.20	0.76	0.65	0.18	0.76	0.65	0.20	0.76
No College	-0.49	-0.86	-0.31	-0.49	-0.84	-0.32	-0.48	-0.85	-0.29

The table presents simulated estimates of mean knowledge by college tier under the status quo, No SAT, and SAT-Optional policies. The SAT-Optional Policy allows students who have not taken the SAT to apply to elite public and private colleges (tiers one and four). Simulated moments are computed using 200 simulated data sets. WA refers to the population of white and Asian students. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

similar, or even more knowledgeable, students.

Going SAT-Optional raises the matriculation of URM students, who are disproportionately unlikely to take the exam, to private elite colleges. Public elite colleges, however, do not see an increase in enrollment, because of the sharp increase in their URM admission threshold. SAT-Optional admissions helps elite colleges enroll more low-income students and the average household income of students attending these colleges falls, from \$105,000 to \$100,000 at elite private colleges and from \$99,000 to \$94,000 at elite public schools. Enrollment at schools that still require the exam falls for low-income and URM students, who become less likely to take the SAT.

7.4 Hypothetical Scenarios

Model estimates have demonstrated the existence of differences in measurement parameters and the mean of initial skills by demographic groups. In this section, I explore how patterns of attendance might be affected if some external agent could alter these parameters. These hypothetical scenarios clarify how patterns of college enrollment are determined not simply by admissions policies in isolation, but by how they interact with the distribution of student skills and the way in which skills map into grades and test scores.

Table 18: Admissions Thresholds, SAT-Optional

	Status Quo		SAT-Optional	
	White/Asian	URM	White/Asian	URM
Tier 1	1.00	0.42	1.09	0.52
Tier 2	0.10	-0.43	0.07	-0.46
Tier 3	-0.39	-0.88	-0.46	-0.94
Tier 4	0.58	0.15	0.77	0.33
Tier 5	0.10	-0.31	0.07	-0.34
Tier 6	-0.33	-0.72	-0.40	-0.79

The table presents admissions thresholds at each college tier under the status quo and SAT-Optional policies. The SAT-Optional Policy allows students who have not taken the SAT to apply to elite public and private colleges (tiers one and four). Thresholds have been standardized by the mean and sd of $\log K_{i,12}$. Simulated moments are computed using 200 simulated data sets. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

Table 19: URM Enrollment and Household Income

	URM Attendance			Household Income		
	Status Quo	No SAT	SAT-Optional	Status Quo	NoSAT	SAT-Optional
Tier 1	0.017	0.016	0.019	105K	103K	100K
Tier 2	0.057	0.057	0.055	90K	87K	91K
Tier 3	0.047	0.048	0.046	73K	71K	73K
Tier 4	0.038	0.036	0.037	99K	96K	94K
Tier 5	0.043	0.041	0.042	89K	86K	90K
Tier 6	0.108	0.107	0.108	72K	70K	73K

The table presents simulated estimates of the fraction of URMs attending each tier of college and mean household income under the status quo, No SAT, and SAT-Optional policies. The SAT-Optional Policy allows students who have not taken the SAT to apply to elite public and private colleges (tiers one and four). Simulated moments are computed using 200 simulated data sets. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

Table 20: Status Quo with $\mathbb{E}[\log K_{i,9} \mid URM_i = 1]$ Raised by 0.5 sd

	URM Attendance		$\mathbb{E}[\log K_{i,12} \vert URM_i = 1, Attend_i = j]$	
	Original	Hypothetical	Original	Hypothetical
Tier 1	0.017	0.021	0.869	1.017
Tier 2	0.057	0.070	0.212	0.351
Tier 3	0.047	0.059	-0.092	0.035
Tier 4	0.038	0.050	0.626	0.731
Tier 5	0.043	0.058	0.288	0.400
Tier 6	0.108	0.141	0.041	0.214
Any College	0.310	0.400	0.203	0.345

The table displays the fraction of URMs attending each school together with their average knowledge under two scenarios. The first scenario simulates the status quo policy using parameter estimates obtained in estimation, while the second, labeled Hypothetical, raises the initial knowledge for every URM student by 0.5 sd before simulating the status quo policy. The third and fourth columns have been standardized by the population mean and sd of $\log K_{i,12}$. Simulated moments are computed using 200 simulated data sets. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

In the first exercise, I raise the initial level of knowledge in the ninth grade by 0.5 sd for URMs and examine how this would affect college attendance in the status quo.[31] Table 20 shows that URM enrollment would increase at all colleges, with gains of over 30% at all public colleges. The average skill level of URMs at each college increases as well.[32] This exercise size shows that, even holding racial preferences fixed, changes in the distribution of skills can cause changes in access to college for URM students.

The second hypothetical exercise explores how patterns of sorting to college would differ if the SAT were eliminated and grades were 50% *more* informative for URMs than for white and Asian students. The estimated measurement parameters in Table 6 revealed that GPAs were less informative for URM students, which may hinder colleges' attempts to enroll skilled URM students in a world without the SAT. Table 21 displays estimates of sorting by knowledge under this second hypothetical exercise. The estimates suggest that, even in the absence of any change in underlying skills, better signals would enable all colleges, but especially elite colleges, to admit stronger URM candidates. With

[31] I do this by increasing the coefficient on URM in the initial skills distribution in Table 4 from -0.65 to -0.15. I do not adjust the distribution of income, which also puts URMs at a disadvantage.

[32] Knowledge of URMs on campus appears to increase by less than the initial investment of 0.5 sd, but this is due to the fact that estimates in the table are standardized by the population mean and sd of $\log K_{i,12}$, which exceeds one.

Table 21: No SAT Policy with Grades 50% More Informative for URMs

	Original			*Hypothetical Scenario*		
	All	URM	WA	All	URM	WA
$\mathbb{E}[\log K_{i,12}]$						
Tier 1	1.15	0.63	1.27	1.28	0.86	1.38
Tier 2	0.67	0.20	0.79	0.68	0.27	0.79
Tier 3	0.44	-0.03	0.58	0.43	0.03	0.55
Tier 4	0.92	0.47	1.01	0.98	0.61	1.07
Tier 5	0.68	0.23	0.78	0.71	0.34	0.79
Tier 6	0.53	0.08	0.63	0.51	0.17	0.6
No College	-0.49	-0.84	-0.32	-0.50	-0.94	-0.29
Complete College	0.692	0.619	0.709	0.688	0.628	0.703
Hours Study	6.36	6.14	6.45	6.36	6.11	6.46

The table displays average knowledge for college students in a scenario where the SAT is eliminated and grades are 50% more informative for URM students than white and Asian students. $\log K_{i,12}$ has been standardized by its mean and standard deviation. Simulated moments are computed using 200 simulated data sets. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

grades now more informative, average knowledge among URMs at elite private colleges increases by over 0.20 sd. These changes in sorting are significant enough to raise URM college completion by 1 pp, comparable with the SAT-for-All policy analyzed earlier.

One reason why banning the SAT does not raise college enrollment for URMs is that the average knowledge of URM applicants falls as more apply to college (Table 14). This suggests that whether banning the SAT increases access to college depends on whether there is a sizable share of strong URM applicants who do not take the SAT but who would apply to college if an SAT score were no longer required. To investigate this, I simulate a policy that eliminates the SAT, but I raise $\log K_{i,9}$ for every URM student who does not take the exam by 0.75 sd so that their initial knowledge is level with typical white and Asian students who take the SAT. Table 22 shows statistics on URM attendance, knowledge, and college completion when banning the SAT under this hypothetical scenario. The fraction of URMs attending college rises by nearly 7.4 pp, and rates of completion rise at every college. This scenario shows that the effects of banning the SAT depend on how the skills of URM students who do not take the exam compare with those who already take it.

Table 22: No SAT Policy with non SAT-Takers more Skilled than Takers

	Status Quo			*No SAT*		
	Attendance	$\mathbb{E}[\log K_{i,12}]$	Complete	Attendance	$\mathbb{E}[\log K_{i,12}]$	Complete
Tier 1	0.016	0.64	0.669	0.020	0.69	0.682
Tier 2	0.057	0.20	0.686	0.068	0.27	0.713
Tier 3	0.048	-0.03	0.571	0.058	0.10	0.629
Tier 4	0.036	0.47	0.712	0.045	0.54	0.716
Tier 5	0.041	0.23	0.652	0.052	0.32	0.676
Tier 6	0.107	0.08	0.555	0.136	0.24	0.605
Any College	0.305	0.18	0.619	0.379	0.29	0.655

The table displays attendance rates, average knowledge, and college completion rates for URM students at each college tier under the status quo and No-SAT policies under a hypothetical scenario in which $\log K_{i,9}$ for URM non-takers is raised by 0.75 sd. Details regarding the scenario are provided in the text. Simulated moments are computed using 200 simulated data sets. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

8 Conclusion

The main goal of this paper is not to provide a single number summarizing the effect of eliminating the SAT on access to four-year colleges for low-income and URM students. Reasonable scholars may disagree over the assumptions, such as how colleges use the SAT, that are necessary to generate such a number. The main goal of the paper is instead to show that when colleges stop using the SAT, they must focus more on other criteria, and this has immediate consequences for who attends college (holding applications fixed) and further affects who applies to college and how well-prepared they are.

The paper has shown that four mechanisms – the shift towards alternative admissions criteria, endogenous applications, endogenous human capital investment, and optimization by capacity-constrained colleges – are quantitatively important in shaping patterns of sorting to college and that they may work against each other. This paper does not find a large effect of eliminating the SAT on access to college for under-represented minorities or for low-income students. But, apart from this null effect, the model has clarified that predicting the effects of eliminating the SAT depends on the answer to the following question: Does there exist a sizable population of low-income, mostly URM students who do not take the exam but who are better-prepared for college than current applicants and who will apply when the exam is removed? If yes, then eliminating the SAT may increase access for disadvantaged students. But if the answer is no, then the mechanisms of endogenous human capital investment and college optimization work to undo the immediate effect of the policy, and colleges will be unable to generate opportunities for disadvantaged applicants without altering their admissions criteria. Such a move could create real economic costs. Estimates from the model reveal that completion at all schools depends strongly on a student's knowledge at the time of matriculation. Accepting less skilled students may lower rates of completion.

There are several important extensions of this research. I have assumed that colleges have preferences over racial diversity and a single latent variable, which I have referred to as knowledge. Anecdotal evidence suggests that colleges also care about other characteristics like extracurricular skills. Although the ELS 2002 does contain self-reported information on extracurricular participation, these variables do not predict admission to college in the ELS 2002 conditional on grades and SAT scores. If future data sets have high quality measures of extracurricular skill (say, athletic or musical skill), I see an important extension in allowing for a two-dimensional set of latent skills and estimating college preferences for diversity and a weighted combination of these two skills. As the ability to invest in extracurriculars like music and art is correlated with income, a greater reliance on these skills in admissions may provide students from richer households with another way of gaining admission in the absence of the SAT. Allowing for a second skill would also reduce uncertainty about the admissions process.

The paper has shown that, without significant changes in preferences or in the distribution of skills, URMs will continue to be under-represented at four-year colleges. It has also shown, consistent with a large literature on early interventions, that policies that raise the skills of URMs prior to high school can raise college attendance and completion (Campbell et al. 2014, Almond, Currie, and Duque 2018, García et al. 2020). Many universities already invest in after-school tutoring and mentoring programs for underprivileged students in their neighborhood. The University of Southern California has for several decades operated a program, the Neighborhood Academic Initiative, that identifies promising middle school students in south Los Angeles and provides them with additional weekend courses on USC's campus to help make them college-ready. USC has also founded several charter schools in south Los Angeles. Together, these investments have provided the university with a steady stream of qualified applicants from under-served backgrounds. The evidence suggests that colleges have the tools to increase opportunity, but that it requires more than a change in admissions criteria.

References

Agostinelli, Francesco, and Matthew Wiswall. 2020. "Estimating the Technology of Children's Skill Formation." Technical Report, Working Paper.

Akhtari, Mitra, Natalie Bau, and Jean-William P Laliberté. 2020. "Affirmative Action and Pre-college Human Capital." Technical Report, National Bureau of Economic Research.

Almond, Douglas, Janet Currie, and Valentina Duque. 2018. "Childhood Circumstances and Adult Outcomes: Act II." *Journal of Economic Literature* 56 (4): 1360–1446.

Anderson, TW, and Herman Rubin. 1956. "Statistical Inference in Factor Analysis." *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability: Held at the Statistical Laboratory, University of California, December, 1954, July and August, 1955*, Volume 1. Univ of California Press, 111.

Arcidiacono, Peter. 2005. "Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?" *Econometrica* 73 (5): 1477–1524.

Arcidiacono, Peter, Esteban M Aucejo, Hanming Fang, and Kenneth I Spenner. 2011. "Does Affirmative Action Lead to Mismatch? A New Test and Evidence." *Quantitative Economics* 2 (3): 303–333.

Arcidiacono, Peter, V Joseph Hotz, Arnaud Maurel, and Teresa Romano. 2020. "Ex Ante Returns and Occupational Choice." *Journal of Political Economy* 128 (12): 4475–4522.

Arcidiacono, Peter, Josh Kinsler, and Tyler Ransom. 2022a. "Asian American Discrimination in Harvard Admissions." *European Economic Review* 144:104079.

———. 2022b. "Recruit to Reject? Harvard and African American Applicants." *Economics of Education Review* 88:102255.

Bodoh-Creed, Aaron, and Brent Richard Hickman. 2017. "Pre-college Human Capital Investments and Affirmative Action: A Structural Policy Analysis of US College Admissions." *Becker Friedman Institute for Research in Economics Working Paper*, no. 2017-10.

Bond, Timothy N, George Bulman, Xiaoxiao Li, and Jonathan Smith. 2018. "Updating Human Capital Decisions: Evidence from SAT Score Shocks and College Applications." *Journal of Labor Economics* 36 (3): 807–839.

Botelho, Fernando, Ricardo A Madeira, and Marcos A Rangel. 2015. "Racial Discrimination in Grading: Evidence from Brazil." *American Economic Journal: Applied Economics* 7 (4): 37–52.

Bowen, William G, and Derek Bok. 2016. "The Shape of the River." In *The Shape of the River*. Princeton University Press.

Bulman, George. 2015. "The Effect of Access to College Assessments on Enrollment and Attainment." *American Economic Journal: Applied Economics* 7 (4): 1–36.

Cameron, Stephen V, and Christopher Taber. 2004. "Estimation of Educational Borrowing Constraints using Returns to Schooling." *Journal of Political Economy* 112 (1): 132–182.

Campbell, Frances, Gabriella Conti, James J Heckman, Seong Hyeok Moon, Rodrigo Pinto, Elizabeth Pungello, and Yi Pan. 2014. "Early Childhood Investments Substantially Boost Adult Health." *Science* 343 (6178): 1478–1485.

Card, David. 1995. "Using Geographical Variation in College Proximity to Estimate the Return to Schooling." *Aspects of Labor Market Behavior: Essays in Honor of John Vanderkamp*, pp. 201–222.

Carlana, Michela. 2019. "Implicit Stereotypes: Evidence from Teachers' Gender Bias." *The Quarterly Journal of Economics* 134 (3): 1163–1224.

Carneiro, Pedro, and James J Heckman. 2002. "The Evidence on Credit Constraints in Post-Secondary Schooling." *The Economic Journal* 112 (482): 705–734.

Carneiro, Pedro, James J Heckman, and Edward J Vytlacil. 2011. "Estimating Marginal Returns to Education." *American Economic Review* 101 (6): 2754–2781.

Carneiro, Pedro, Michael Lokshin, and Nithin Umapathi. 2017. "Average and Marginal Returns to Upper Secondary Schooling in Indonesia." *Journal of Applied Econometrics* 32 (1): 16–36.

Cotton, Christopher S, Brent R Hickman, and Joseph P Price. 2022. "Affirmative Action and Human Capital Investment: Evidence from a Randomized Field Experiment." *Journal of Labor Economics* 40 (1): 157–185.

Cunha, Flavio, James J Heckman, and Susanne M Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3): 883–931.

Currie, Janet, and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Vapital: Evidence from College Openings." *The Quarterly Journal of Economics* 118 (4): 1495–1532.

Delavande, Adeline, and Basit Zafar. 2019. "University Choice: The Role of Expected

Earnings, Nonpecuniary Outcomes, and Financial Constraints." *Journal of Political Economy* 127 (5): 2343–2393.

Dillon, Eleanor Wiske, and Jeffrey Andrew Smith. 2017. "Determinants of the Match between Student Ability and College Quality." *Journal of Labor Economics* 35 (1): 45–66.

Donni, Olivier. 2015. "Measuring the Cost of Children." *IZA World of Labor*.

Dynarski, Susan M. 2018. "ACT/SAT for All: A Cheap, Effective Way to Narrow Income Gaps in College." *Brookings Institution*.

Epple, Dennis, Richard Romano, and Holger Sieg. 2006. "Admission, Tuition, and Financial Aid Policies in the Market for Higher Education." *Econometrica* 74 (4): 885–928.

———. 2008. "Diversity and Affirmative Action in Higher Education." *Journal of Public Economic Theory* 10 (4): 475–501.

Fu, Chao. 2014. "Equilibrium Tuition, Applications, Admissions, and Enrollment in the College Market." *Journal of Political Economy* 122 (2): 225–281.

García, Jorge Luis, James J Heckman, Duncan Ermini Leaf, and María José Prados. 2020. "Quantifying the Life-cycle Benefits of an Influential Early-childhood Program." *Journal of Political Economy* 128 (7): 2502–2541.

Golightly, Eleanor. 2019. "Does College Access Increase High School Effort? Evaluating the Impact of the Texas Top 10% Rule." Technical Report.

Goodman, Joshua, Oded Gurantz, and Jonathan Smith. 2020. "Take Two! SAT Retaking and College Enrollment Gaps." *American Economic Journal: Economic Policy* 12 (2): 115–58.

Goodman, Sarena. 2016. "Learning from the Test: Raising Selective College Enrollment by Providing Information." *Review of Economics and Statistics* 98 (4): 671–684.

Grau, Nicolás. 2018. "The Impact of College Admissions Policies on the Academic Effort of High School Students." *Economics of Education Review* 65:58–92.

Hastings, Justine, Christopher A Neilson, and Seth D Zimmerman. 2015. "The Effects of Earnings Disclosure on College Enrollment Decisions." Technical Report, National Bureau of Economic Research.

Hoxby, Caroline M, and Christopher Avery. 2012. "The Missing "One-offs": The Hidden Supply of High-Achieving, Low Income Students." Technical Report, National Bureau of Economic Research.

Hyman, Joshua. 2017. "ACT for All: The Effect of Mandatory College Entrance Exams on Postsecondary Attainment and Choice." *Education Finance and Policy* 12 (3): 281–311.

Kane, Thomas J, and Cecilia Elena Rouse. 1995. "Labor-Market Returns to Two- and Four-year College." *The American Economic Review* 85 (3): 600–614.

Kapor, Adam. 2020. "Distributional effects of race-blind affirmative action." Technical Report, Working Paper.

Klasik, Daniel. 2013. "The ACT of Enrollment: The College Enrollment Effects of State-required College Entrance Exam Testing." *Educational Researcher* 42 (3): 151–160.

Kling, Jeffrey R. 2001. "Interpreting Instrumental Variables Estimates of the Returns to Schooling." *Journal of Business & Economic Statistics* 19 (3): 358–364.

Leeds, Daniel M, Isaac McFarlin, and Lindsay Daugherty. 2017. "Does Student Effort Respond to Incentives? Evidence from a Guaranteed College Admissions Program." *Research in Higher Education* 58 (3): 231–243.

Otero, Sebastián, Nano Barahona, and Cauê Dobbin. 2021. "Affirmative Action in Centralized College Admission Systems: Evidence from Brazil." Technical Report, Working paper.

Pallais, Amanda. 2015. "Small Differences that Matter: Mistakes in Applying to College." *Journal of Labor Economics* 33 (2): 493–520.

Rauschenberg, Samuel. 2014. "How Consistent Are Course Grades? An Examination of Differential Grading." *Education Policy Analysis Archives* 22 (92): n92.

Schmill, Stu. 2022. We Are Reinstating Our SAT/ACT Requirement for Future Admissions Cycles.

Soares, Joseph A. 2020. "The Scandal of Standardized Tests: Why We Need to Drop the SAT and ACT." *Teachers College Press*.

Tincani, Michela M, Fabian Kosse, and Enrico Miglino. 2021. "Subjective Beliefs and Inclusion Policies: Evidence from College Admissions." Technical Report, Working Paper.

Tough, Paul. 2019. "What College Admissions Offices Really Want." *The New York Times*.

Van den Bergh, Linda, Eddie Denessen, Lisette Hornstra, Marinus Voeten, and Rob W Holland. 2010. "The Implicit Prejudiced Attitudes of Teachers: Relations to Teacher Expectations and the Ethnic Achievement Gap." *American Educational Research Journal* 47 (2): 497–527.

Williams, Benjamin. 2020. "Identification of the Linear Factor Model." *Econometric Reviews* 39 (1): 92–109.

Wiswall, Matthew, and Basit Zafar. 2015. "Determinants of College Major Choice: Identification Using an Information Experiment." *The Review of Economic Studies* 82 (2): 791–824.

Appendices

A Sample Selection Criteria

Of the 16,200 students who were initially sampled in the ELS 2002, 12,880 responded in both the base year and the first follow-up survey in 2004. The most common reasons for exclusion are nonresponse (12.05 % of the sample), dropping out of high school between 2002 and 2004 (3.76% of the sample), and graduating early (2.43% of the sample). Other, infrequent, reasons include being out of the country, language difficulties making the survey impossible, and death. 370 students lack information on the amount of time they spend studying. A further 2,370 students either lack information on GPA, school characteristics, or geocode data. 610 individuals lack an SAT score despite applying to colleges that required the exam. Finally, I exclude a small number of individuals (< 10) with extremely low grades and SAT scores who are admitted to elite colleges, possibly because of athletics. These students cause the model likelihood function to return infinite values for large regions of the parameter space. These exclusions result in a sample of 9,910 observations.

During the second follow-up survey in 2006, students report the full list of colleges they applied to, where they were admitted, and where they first matriculated. To address concerns regarding whether use of this self-reported measure may result in biased admissions probabilities, I compare admission probabilities derived from the survey responses in the ELS 2002 to the official probabilities in the Integrated Postsecondary Education Data System (IPEDS) for the same year (2004/05 cohort) in Table A-1. The table disaggregates admission probabilities by type (private non-profit vs public) and Barron's Selectivity Ranking.[33] The IPEDs statistics are weighted by the enrollment of the institution. The table shows that admissions rates are somewhat lower in IPEDS than in the ELS 2002. The differences, of eleven to thirteen percentage points (pp) for private colleges, and between six and eight pp for public colleges, indicate that students in the ELS 2002 selectively under-report applications to colleges at which they are rejected. The implication for the empirical results is that the cost of applying to college, which is identified by the number of applications to college, could be over-estimated in the model. This suggests that model forecasts of application growth resulting from policy changes are likely to be a lower bound for the true application growth.

[33] This grouping of colleges into six tiers will be maintained throughout the paper.

Table A-1: Admission Rates, IPEDS and ELS 2002

			Rate of Admission	
Type	Tier	Barron's Rank	IPEDS	ELS 2002
Private	1	1	0.29	0.40
	2	2 and 3	0.66	0.79
	3	4, 5 and 6	0.72	0.84
Public	4	1 and 2	0.53	0.61
	5	3	0.69	0.75
	6	4, 5, and 6	0.72	0.80

The table compares rates of admission at six college tiers based on self-reported application and admission information in the ELS 2002 together with data from IPEDs for the same set of schools. Admission rates for each school are weighted by enrollment to compute the IPEDs statistics. SOURCES: IPEDS and National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

B Alternative Normalizations and Specifications for the Dynamic Factor Model

The identification of dynamic factor models requires normalizations. In this paper, I exploit the fact that the tenth and twelfth grade NCES math exams are scored according to item response theory to reduce the number of necessary normalizations. However, as I explain in section 4.1, since both the parameters of the measurement system and the mean of the initial conditions vary by whether a student belongs to an under-represented minority, an additional normalization is required. The approach I adopted in this paper is to impose that the exams scored by item response theory have the same constants in equation 2, namely that $\mu_{10,j}^{URM} = \mu_{10,j}^{WA}$ and $\mu_{12,j}^{URM} = \mu_{12,j}^{WA}$ for j equal to the NCES math exam. In this section, I explore whether the inferences I draw from the dynamic factor model are robust to alternative normalizations and alternative specifications for the technology of skill formation.

Tables B-1 and B-2 present estimates from a dynamic factor model that instead imposes the normalization that GPA in the ninth grade has the same constant for URM students as it does for white and Asian students, $\mu_{9,j}^{URM} = \mu_{9,j}^{WA}$ for j equal to the ninth grade GPA. The NCES math exams are now permitted to have different constants by URM status. The estimates of bias in Table B-1 are qualitatively and quantitavely very similar to the main specification. Despite using a different normalization, I estimate that neither of the SAT exams, nor any of the NCES exams are biased. Similar to the main specification, the only evidence of bias is in GPA in the twelfth grade, which appears to be biased against URMs. The estimates of signal-to-noise ratios in table B-2 are qualitatively very similar to those from the main specification. I estimate that GPAs are less informative for URMs during each year of high school, and I find that the standardized exams are typically more informative for URMs than for white and Asian students.

The model in the main part of the paper specificies a deterministic skill technology (equation 1). I now explore whether the results are robust to the inclusion of a shock in this equation. Because of the inclusion of the shock, the identification argument in section 4.1 breaks down, and additional normalizations are needed. Since there is only one measurement in grade nine, it is not possible to separately identify the variance of the technology shock, the variance of the measurement shock, and the factor loading, so I normalize the factor loading on GPA in the ninth grade to equal the factor loading on GPA in the tenth grade, $\alpha_{9,j}^{R} = \alpha_{10,j}^{R}$ for $j = GPA$ and $R = URM, WA$. For the same reason, I normalize the factor loadings on eleventh and twelfth grade GPAs to be the same, $\alpha_{11,j}^{R} = \alpha_{12,j}^{R}$, again for $j = GPA$ and $R = URM, WA$.

Table B-3 presents estimates from this dynamic factor model. The estimated variance of the shock is not statistically different from zero, and the rest of the parameters are nearly identical to the main specification in the paper (Table 7). The results presented here should alleviate concern that the use of a deterministic skill formation equation, which aids in identification, has significant effects on the results.

Table B-1: Bias

	URM	White & Asian	Difference
GPA, 9th grade	-0.17	-0.17	0
	(0.03)	(0.03)	(NA)
GPA, 10th grade	-0.18	-0.16	-0.02
	(0.03)	(0.03)	(0.03)
GPA, 11th grade 11	-0.20	-0.15	-0.05
	(0.03)	(0.02)	(0.03)
GPA, 12th grade 12	-0.26	-0.10	-0.15
	(0.03)	(0.02)	(0.05)
SAT Math	-0.66	-0.69	0.03
	(0.04)	(0.03)	(0.03)
SAT Verbal	-0.59	-0.6	0.01
	(0.04)	(0.03)	(0.03)
NCES Reading, 10th grade	-0.17	-0.20	0.03
	(0.04)	(0.03)	(0.03)
NCES Math, 10th grade	-0.22	-0.23	0.02
	(0.04)	(0.03)	(0.03)
NCES Math, 12th grade	-0.22	-0.23	0.02
	(0.04)	(0.03)	(0.03)

The table displays estimates of $\boldsymbol{\mu}_t^{URM}$ in equation 1 when ninth grade GPA is normalized to have no bias. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

Table B-2: Signal-to-Noise Ratios

	URM	White & Asian	Difference
GPA, 9th grade	0.77	0.93	-0.15
	(0.03)	(0.03)	(NA)
GPA, 10th grade	0.77	0.92	-0.16
	(0.04)	(0.03)	(0.03)
GPA, 11th grade	0.69	0.8	-0.12
	(0.03)	(0.03)	(0.03)
GPA, 12th grade	0.55	0.6	-0.05
	(0.02)	(0.02)	(0.03)
SAT Math	1.87	1.87	0.00
	(0.09)	(0.06)	(0.08)
SAT Verbal	1.47	1.25	0.22
	(0.09)	(0.04)	(0.08)
NCES Reading, 10th grade	1.28	1.04	0.23
	(0.06)	(0.03)	(0.05)
NCES Math, 10th grade	2.07	1.65	0.42
	(0.11)	(0.05)	(0.01)
NCES Math, 12th grade	2.34	1.92	0.41
	(0.09)	(0.06)	(0.06)

The table displays estimates of signal-to-noise ratios in equation 2 when ninth grade GPA is normalized to have no bias. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

Table B-3: Technology of Skill Formation

	URM	White & Asian
Knowledge(-1)	0.98	1.01
	(0.04)	(0.01)
Study, 10 hours/wk	0.10	0.08
	(0.03)	(0.01)
Private School	0.05	0.01
	(0.04)	(0.02)
Free Lunch	-0.14	-0.19
	(0.00)	(0.01)
Student Teacher Ratio	0.07	-0.06
	(0.06)	(0.07)
Mother: High School	-0.02	-0.02
	(0.01)	(0.01)
Mother: Some College	-0.04	-0.03
	(0.02)	(0.02)
Mother: Bachelors	-0.04	-0.04
	(0.06)	(0.07)
Mother: Postgraduate	-0.04	-0.04
	(0.01)	(0.01)
σ_k^2	0.00	0.00
	(0.40)	(0.40)

The table displays estimates of parameters the technology of skill formation with a stochastic shock. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

C First Stage Regressions

Table C-1: Effect of Testing Availability and Income

	Take SAT			*Apply to College*		
Testing Dates	0.055 (0.016)	0.056 (0.016)	0.042 (0.017)	0.057 (0.015)	0.059 (0.015)	0.047 (0.015)
School Type	Yes	Yes	Yes	Yes	Yes	Yes
School Size	Yes	Yes	Yes	Yes	Yes	Yes
Local Poverty Rate	No	Yes	Yes	No	Yes	Yes
Mother's Education	No	No	Yes	No	No	Yes
N	9910	9910	9910	9910	9910	9910

The table displays the results of logit regressions of binary indicators for taking the SAT and applying to college on the number of SAT testing dates in one's own school during spring of the junior year of high school. Controls for the type of school, enrollment, district level poverty rates, and mother's education (5 categories) are also included. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

Table C-2: Effect of Distance on Applications

	Apply to College					
	Private			*Public*		
	Tier 1	Tier 2	Tier 3	Tier 4	Tier 5	Tier 6
Distance : 50-150 miles	-0.0121***	-0.0156***	-0.0122***	-0.0479***	-0.0686***	-0.0645***
	(0.0016)	(0.0009)	(0.0006)	(0.0042)	(0.0039	(0.0020)
Distance : 150 - 250 miles	-0.0155***	-0.0203***	-0.0145***	-0.0735***	-0.0942***	-0.0773***
	(0.0016)	(0.0009)	(0.0006)	(0.0042)	(0.0039)	(0.0020)
Distance : 250 - 500 miles	-0.0179***	-0.0221***	-0.0152***	-0.0800***	-0.1034***	-0.0807***
	(0.0016)	(0.0009)	(0.0006)	(0.0042)	(0.0039	(0.0020)
Distance : over 500 miles	-0.0199***	-0.0230***	-0.0153***	-0.0893***	-0.1056***	-0.0817***
	(0.0016)	(0.0009)	(0.0006)	(0.0043)	(0.0039)	(0.0020)
N	9910	9910	9910	9910	9910	9910
Num Schools	60	270	480	40	80	380
Individual FEs	Yes	Yes	Yes	Yes	Yes	Yes
Clustered SEs	Individual	Individual	Individual	Individual	Individual	Individual

The table shows the coefficients from linear probability models of application decisions. The unit of observation is an individual-college. All four-year public and private non-profit colleges are included. Distance is calculated using the individual's home census block while in high school and the latitude and longitude of each school in IPEDs. Each regression includes individual fixed effects and controls for tuition, acceptance rates, and interactions between acceptance rates and an individual's performance on the 12th grade NCES math exam. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

D Knowledge Predicts Admission

This paper has modeled each student's probability of admission to college as a function of her demographic and her knowledge at the time of application, $\log K_{i,12}$, as filtered through the observable measurements in $\Omega_{i,12}$. This section shows that knowledge filtered in this way is highly predictive of admission.

Figures D-1 through D-6 provide estimates of nonparametric regressions of admission to college as a function of the mean of each student's twelfth grade knowledge as derived in the model, $\mathbb{E}[\log K_{i,12} \mid \Omega_{i,12}]$. The regressions are all estimated using an Epanechnikov kernel and a bandwidth of one. The nonparametric functions are plotted only over the support of mean knowledge among applicants to each school in the ELS 2002. The figures show that admission probabilities are increasing in the mean of knowledge at all schools. While many colleges observe characteristics that are not in the ELS 2002, such as writing samples and teacher recommendations, the nonparametric regressions provide reassurance that the measurements observed in the data are still highly predictive of admission.

Figure D-1: Probability of Admission, Elite Private Colleges

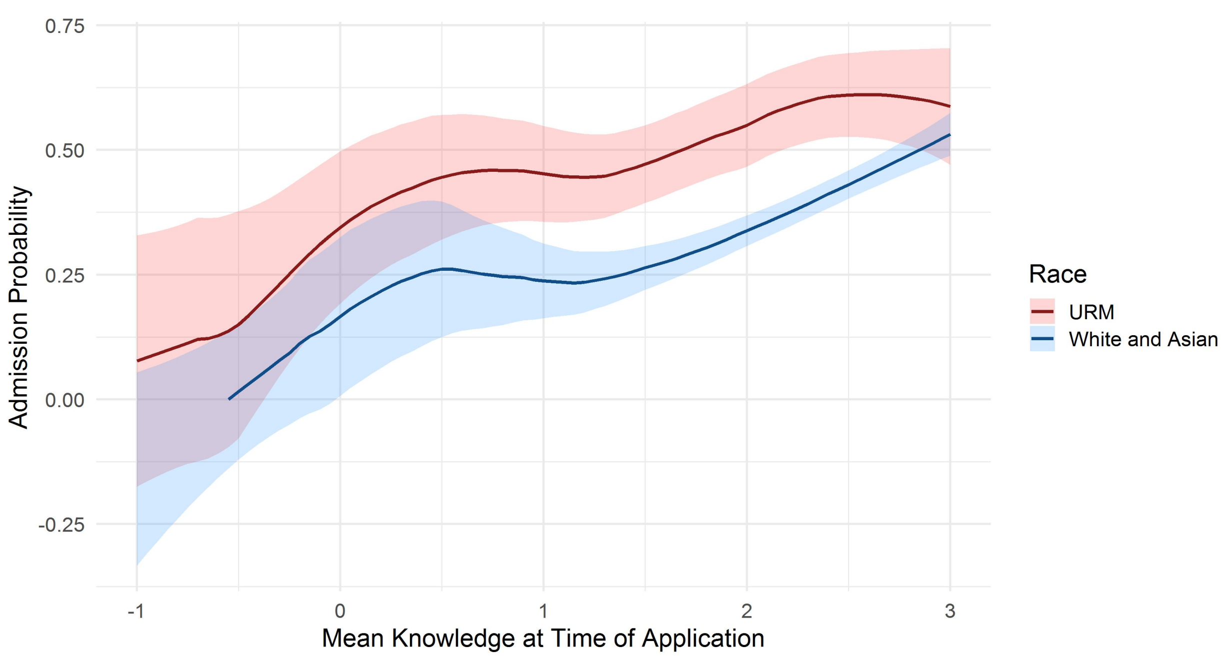


Figure D-2: Probability of Admission, Highly Selective Private Colleges

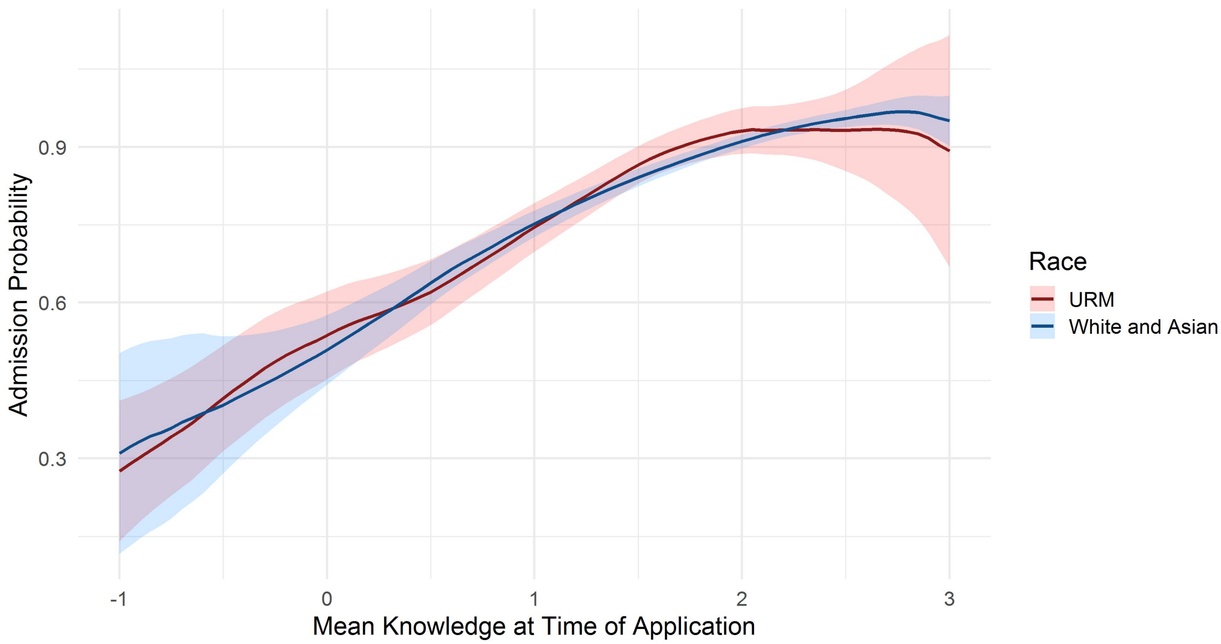


Figure D-3: Probability of Admission, Less Selective Private Colleges

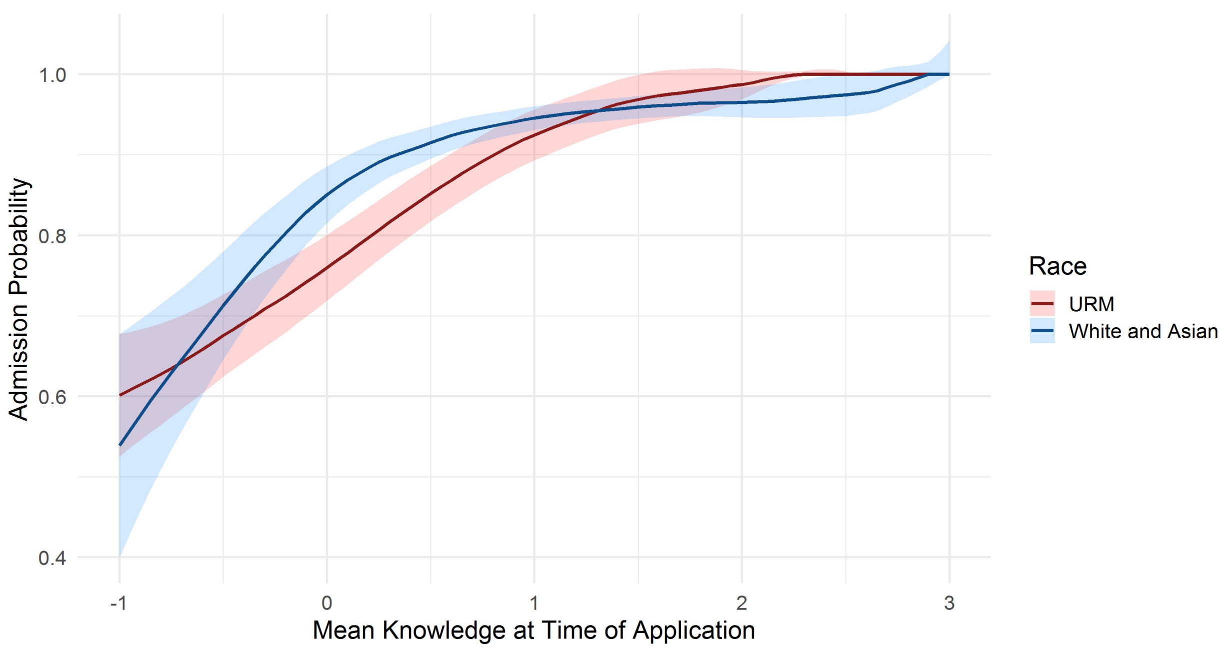


Figure D-4: Probability of Admission, Elite Public Colleges

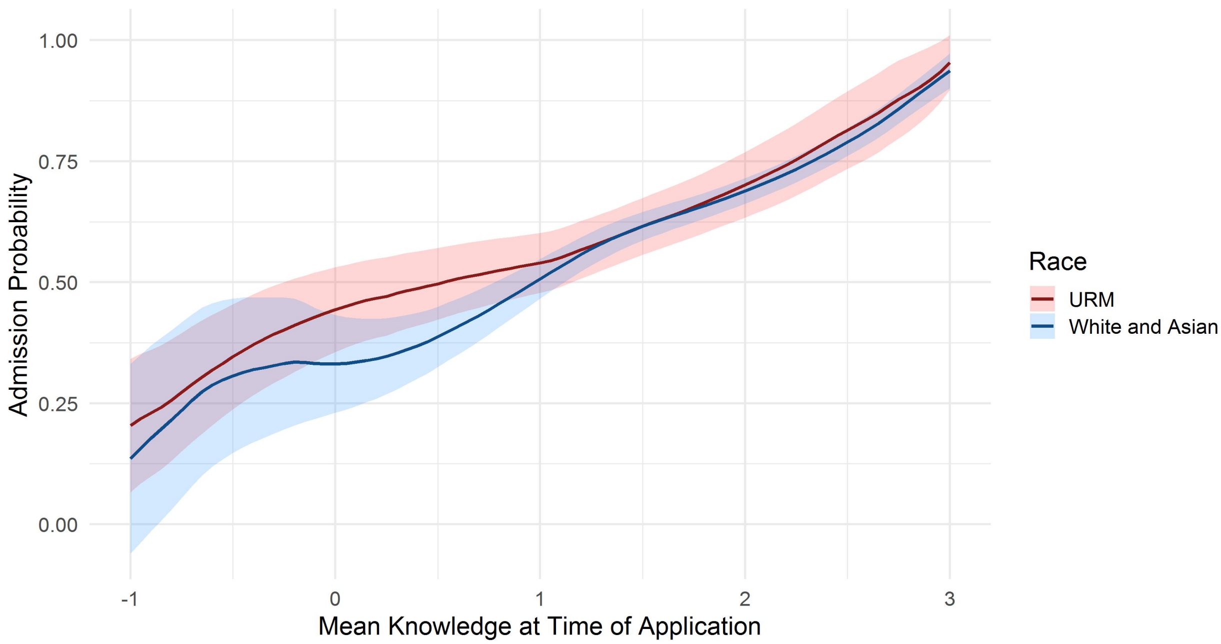


Figure D-5: Probability of Admission, Typical State Flagships

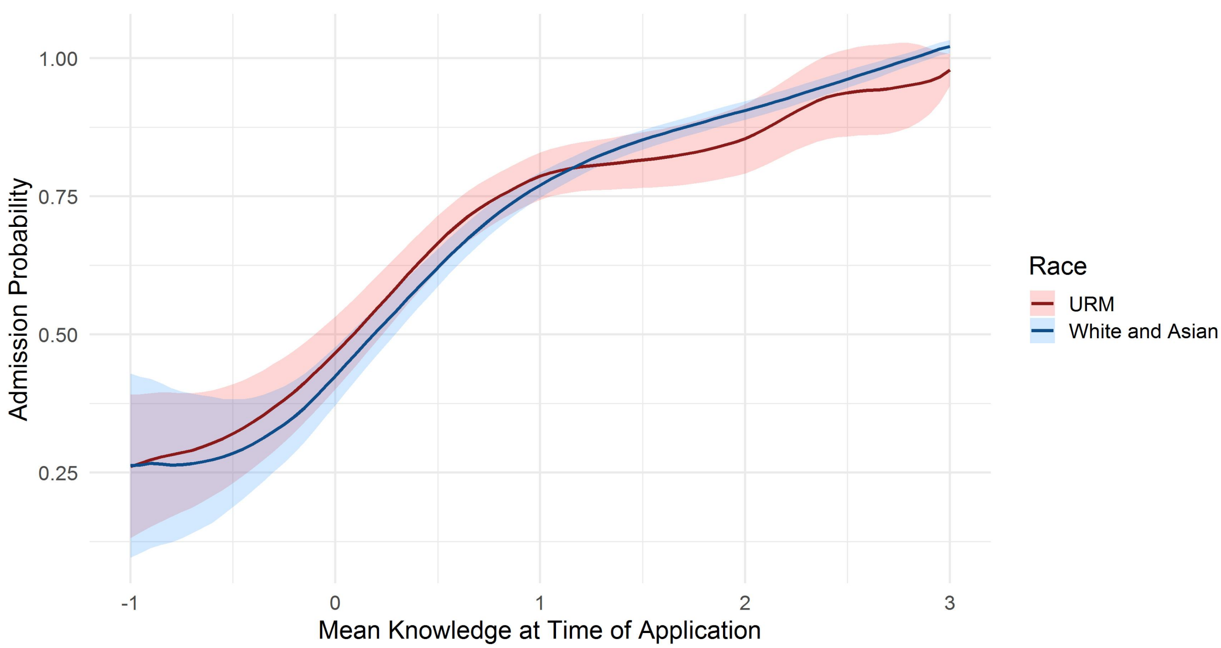


Figure D-6: Probability of Admission, Typical State Satellites